A PHARMACEUTICAL INDUSTRIAL MEDICAL LIBRARY: INDEXING AND INFORMATION RETRIEVAL

Jennifer Elizabeth de Wet

A Dissertation Submitted to the Faculty of Arts
University of the Witwatersrand, Johannesburg
for the Degree of Master of Arts

Johannesburg 1978

I declare that this is my own unaided work with the exception of the statistical analyses in Chapter 4 which were performed for me on a Hewlett Packard series 9800 model 30 minicomputer by Mrs. L. Salter.

No part of this work has been submitted in the past for a degree in any university.

The information used in this dissertation was obtained by me while employed by Ciba-Geigy (Pty) Ltd.

CONTENTS

PAGE

I declare that this is my own unaided work with the exception of the statistical analyses in Chapter 4 which were performed for me on a Hewlett Packard series 9800 model 30 minicomputer by Mrs. L. Salter.

No part of this work has been submitted in the past for a degree in any university.

The information used in this dissertation was obtained by me while employed by Ciba-Geigy (Pty) Ltd.

ABSTRACT

The generation, acquisition, utilization and transfer of information are examined in a medical library in the pharmaceutical industry.

The library system is evaluated from both a user and a system management viewpoint. User information requirements were determined and the dissemination of information from the library was studied.

The study shows the results of research into the use of the library and establishes which library system components facilitated user satisfaction and which were ineffective in providing access to stored information.

The emphasis on performance requirements from the system resulted in the experimental evaluation of indexes in use in the library system as well as in an evaluation of the system's overall operating efficiency during the retrospective searching process. Current retrospective searching techniques proved effective.

The growth of the library as a result of user demand, together with the expressed demand from the medical and pharmacy professions for pharmacological information indicates that there is an increasing demand for information services in this subject field. Possibilities for library co-operation are examined.

ACKNOWLEDGEMENTS

My sincere thanks and appreciation to Mrs. Ursula Brigish, and Professor R. Charlton my supervisors, for their guidance, and helpful suggestions.

I would like to thank Dr. Dudley Jacobs for allowing me to obtain the data for this study while employed in his department in Ciba-Geigy (Pty) Ltd.

My thanks to Mrs. J. Appleton for typing the manuscript and to Miss M. Hall for the assistance rendered in gathering data. I appreciate the assistance of the Department of Library Science staff of the Randse Afrikaanse Universiteit and the University of the Witwatersrand who made possible the index experiment. To the students themselves go my thanks for the time and effort which they put into the experiment.

In addition, I should like to thank Mrs. Lesley Salter for her programming of some of the statistical data.

INTRODUCTION

1. PURPOSE OF THE STUDY

This study examines four aspects of information - its generation, acquisition, utilization and transfer within a medical library in the pharmaceutical industry. The library operates in a commercial and industrial envi onment, covering a restricted subject field for users employed by the company, as well as users from the medical and pharmacy professions and research workers in related occupations. The emphasis is placed upon the discovery of the patterns of information use and dissemination, as well as the information retrieval needs of the actual users of the library service.

Measurement criteria for the evaluation of the library system are discussed and certain criteria are applied in order to evaluate the system a) from a user point of view in order to facilitate planning to meet future user requirements;

b) from a system management viewpoint in order to ensure the continual evolution and adaptation of the service to meet changing information demands made of a special library in a dynamic and developing industrial environment.

A great drawback in any study undertaken in the pharmaceutical industry is the fact that it is impossible to cover the industry as a whole, because of strong competition amongst companies leading to stringent industrial secrecy, and any study must of necessity, be restricted to one such company. For this reason, no inter-company comparative study of drug information storage, retrieval and dissemination was possible.

The purpose of this study is, therefore, to review, analyse and evaluate a library system within a single company in the pharmaceutical industry and to evaluate the library system's performance in respect of the retrieval and dissemination of information from the system, as well as to estimate potential user information requirements from outside the industrial environment in which the library operates. The study shows results of research into the use of the system and the evaluation indicates the operating effectiveness of the system as a whole, as well as selected subcomponents of the system, viz. the indexes.

The use of the library system depends only to a small extent on its efficiency as evaluated, but depends rather to a larger extent on whether the system characteristics correspond with the situation and requirements of the user.

It is for this reason that an analysis of expressed information needs was undertaken. The considerable gap between information supply and demand has been investigated chiefly from its 'objective' aspect, namely the problems of the rapidly increasing bulk of scientific and technical data which overloads conventional ways of information transmission and encourages the use of new techniques. The 'subjective' aspect of this information system - the user - has not, however, been relegated to a position of minor importance, but has been examined in some detail in an attempt to establish some pattern in the information disseminated from the library.

This study is, therefore, firstly a user study, then a dissemination study and lastly a use study. In other words, this is a study of user demand i.e. the information needs expressed by the user groups

The purpose of this study is, therefore, to review, analyse and evaluate a library system within a single company in the pharmaceutical industry and to evaluate the library system's performance in respect of the retrieval and dissemination of information from the system, as well as to estimate potential user information requirements from outside the industrial environment in which the library operates. The study shows results of research into the use of the system and the evaluation indicates the operating effectiveness of the system as a whole, as well as selected subcomponents of the system, viz. the indexes.

The use of the library system depends only to a small extent on its efficiency as evaluated, but depends rather to a larger extent on whether the system characteristics correspond with the situation and requirements of the user.

It is for this reason that an analysis of expressed information needs was undertaken. The considerable gap between information supply and demand has been investigated chiefly from its 'objective' aspect, namely the problems of the rapidly increasing bulk of scientific and technical data which overloads conventional ways of information transmission and encourages the use of new techniques. The 'subjective' aspect of this information system - the user - has not, however, been relegated to a position of minor importance, but has been examined in some detail in an attempt to establish some pattern in the information disseminated from the library.

This study is, therefore, firstly a user study, then a dissemination study and lastly a use study. In other words, this is a study of user demand i.e. the information needs expressed by the user groups

concerned with the communication process and mechanisms of scientific and technical information within the pharmaceutical industry in South Africa i.e. towards the medical and allied professions. In addition, this is a use study and attempts to characterize information habits and behaviour in this specific environment in the context of the organizational and occupational pattern of the specific user groups. Actual users of the system have been studied by means of a user survey. The questioning techniqe utilized was the oral interview which is the most frequently used method in conventional empirical user analysis. Interviews have the advantage of being more precise, spontaneous, flexible and adaptive to spontaneous situations, especially in this environment where little is known about user demand so that a standardized questionnaire could not be designed.

In this study, reference question/inquiries were analysed with regard to their content and were grouped into three categories with the following percentages: simple data 47%; more detailed information 39%; comprehensive detailed information 13%.

It was the objective of this investigation to increase the effectiveness of this information system by investigating the information habits, behaviour and needs of its users, by eliciting user satisfaction and finally by attempting to discover to what extent those requirements were met by the system. It was intended to establish which system attributes facilitated user satisfaction and which did not.

Due to the information explosion and the increasing costs of library services, the more efficient organization of information facilities has become more urgent. The user of these facilities should be as satisfied as possible, given the available resources. Information

facilities should be seen as components of the communication process and the user of these as elements of these systems. Perf rmance effectiveness of the system is the extent to which the system satisfies user demand in respect of relevance to subject information field; kinds of service offered; forms of literature provision; quality of service; timeliness of information; speed of information transfer; recall and precision; comprehensiveness of information coverage; and ease of access to the information stored. By defining the overall performance of the services offered as well as the structure of the system, with emphasis on the above aspects, it is possible to measure and improve the efficiency of the system.

This emphasis on performance requirements of the system resulted in the further evaluation of certain system components, namely the indexes in use in the system. The experiemental evaluation of Index Medicus together with two indexes in use in the system, revealed problem areas in indexing language in one of the system indexes.

A further study of total system operating efficiency was then undertaken. This indicated that the current retrospective searching techniques were effective. Because of the poor test results of the one company index in the index experiment, a further analysis of indexing language in both indexes in use in the system was embarked upon. An independent judge found that both indexes had a high proportion of correctly assigned indexing terms in common, whereas the one which had performed poorly in the index experiment had a considerable number of incorrectly assigned terms. For this reason, a more detailed comparative analysis of indexing language was included.

2. SCOPE

The scope of the study is limited to a study of the library's actual and expected users during a specific period and to factors which influence user satisfaction, as well as tc indexing and retrospective searching methods which indicate the performance achievements of the service.

The study covers the period 1972 - 1978 in Ciba-Geigy (Pty) Limited, South Africa, and, within this context, provides an indication of the effectiveness of a partially mechanized information retrieval system.

3. DEFINITION OF TERMS

It is necessary to include definitions of certain key terms - which may differ from definitions accepted by others - in order to illustrate the intended meaning.

3.1 Adverse Drug Reactions (ADR)

"The mechanisms of adverse drug reactions are threefold) i) intolerance - sensitivity with lack of capacity to endure; ii) idiosyncrasy - susceptibility peculiar to an individual and which causes that individual to react differently to most persons; iii) hypersensitivity - a condition of abnormal sensitivity in which the response to a stimulus is unusually prompt or excessive". [1]

3.2 Communication

3.2.1 "Communication is the control of behaviour through descriptive and reinforcing stimuli". [2]

3.2.2 "... any occurence involving a minimum of four sequential ingredients; a generator of a sign-symbol system which is projected to at least one receiver who assigns meaning". [3]

3.2.3 "Communication has as its central interest those behavioural situations in which a source transmits a message to (a) receiver(s) with conscious in-

---

1 JACOBS, H.D.: Clinical evaluation of side effects of drugs. In: South African medical Journal 47(30): 1365, 1973.

2 HARTMAN, Frank: quoted in Andersch, E.G.: Communication in everyday use. 3rd ed. New York, Holt, Rinehart and Winston, 1969. p. 14.

3 GOYER, Robert: op.cit. p. 15.

tent to affect the latter's behaviours".[4]

3.3 A Current awareness service is a service instituted by an information centre which undertakes to keep its expected immediate clientele (i.e. the expected immediate potential user group) regularly informed of new publications in the subject fields covered by the centre. The service frequently takes the form of an abstract bulletin, but may even be an accession list of new acquisitions. It differs from an S.D.I. service in that is is instituted for the benefit of the group which the centre serves as distinct from the S.D.I. service which provides individuals in the group with new data pertaining to their specific stated interests.

3.4 Dissemination of information

3.4.1 The operation of transferring information from its store to its users.

3.4.2. The professional communication process with regard to sources, authors and mechanisms of transfer of that information.

3.5 Drugs are substances with palliate or cure disease or prevent illness or maintain health.

3.6 Information demand is the total of all information needs of a distinct discipline of user group.

3.7 Information flow is the way in which information is transferred from the source, i.e. the generator, to the utilizer.

---

4 MILLER, Gerald R.: On defining communication: another stab. In: Journal of communication 16: 92, 1966.

3.8 Information needs are the individual needs of users.

3.9 Information requirements are imposed upon the individual user through the nature of his task and his position in the organizatien surrounding him. The user must meet these requirements if he wants to accomplish his task.

3.10 Information retrieval systems are "systems that retrieve documents or document references" [5] - they do not inform (i.e. change the knowledge of) the user on the subject of his inquiry. They "inform him on the existence (or nonexistence) and whereabouts of documents relating to his request". [6]

"An information retrieval system may retrieve complete texts of documents, document surrogates (such as abstracts), or names and addressed of documents (i.e., full bibliographic citations)". [7]

An IR system is a complex phenomenon embracing documents, requests, shorthand descriptions of these documents and requests, a mechanism to allow matching of these descriptions, and people. The people involved are of two types: (a) information staff, or system operators, who describe documents input to the system (i.e., index them), maintain a file of

---

5 LANCASTER, F. Wilfrid: op.cit. p. 2.

6 Ibid. p. 1.

7 Loc.cit.

these descriptions (i.e., an index), and search the file; and (b) system users, who put requests to the file". [8]

3.11 Library efficacy (effectiveness). The efficacy of a library system is the extent to which the goals of the system are achieved. Because of the library's service function, therefore, efficacy must be viewed from two viewpoints - that of the operators of the system and that of the system users. (The latter viewpoint has been adopted in this study).

3.12 Library efficiency is measured by the use of the recall and precision ratios and is an indication of the performance of the library i.e. its operating efficiency - its performance in relation to user needs.

3.13 Pharmacology is "the scientific discipline concerned with the interactions of chemical agents (drugs) and living material, whether or not the actions are good or bad for the living material, or whether the living material is plant or animal in origin". [9]

3.14 The Precision ratio is the number of relevant documents retrieved, in proportion to the total number of documents retrieved.

3.15 The Recall ratio is the number of relevant documents retrieved in proportion to the total number of relevant documents in the collection.

---

8 Ibid. p. 2.

9 LEAKE, Chauncey D.: An historical account of pharmacology to the 20th century.
Springfield, Ill., Charles C. Thomas, (1975). p. 3.

3.16 Regression analysis (simple linear)

By the use of regression analysis, scores on the dependent variable can be predicted from the independent variable.

x = independent variable

y = dependent variable

The y variable i.e. the dependent variable is so called since any y value depends on the population sampled.

3.17 Scientific and technical information is defined as "any information which is required from time to time to keep abreast of new technological developments, to develop even further existing products or processes, to maintain equipment, to solve production problems or which is needed for the establishment of new factories". [10]

3.18 Special library. The American Special Libraries Association defines a special library as follows:

3.18.1 a) "A library or information centre maintained by an individual, corporation, association, government agency or any other group; or

b) A specialized or departmental collection within a library:

For the organization and dissemination of information, and primarily offering service to a specialized clientele through the use of various media and methods". [11]

---

10 F C I viewpoint 8(3): 9, 1974.

11 Special libraries 67(10): 78, 1976.

3.18.2 Bhattacharyya defines special libraries as "those centres that are subject-oriented and designed to serve 'not all readers' but only those with specified interest". [12]

3.19 Selective dissemination of information (SDI) service is "a service in which the interests of an individual, or group of individuals, are described by means of words, or phrases (i.e., a 'user interest profile' is created). When documents input to a collection are similarly described in an indexing operation, the document profiles thus created may be matched against the file of user interest profiles. When a match of profiles above a certain pre-established threshold occurs, the user is informed of the existence of this document, which is presumed to be relevant to his interests". [13]

3.20 Toxicology. "When one studies the harmful effects of drugs or chemicals on living things, one is concerned with toxicology, from the Greek toxicon, a poison". [14]

"Practically pharmacology and toxicology go together, for the harmful or poisonous or toxic effects of drugs are usually the result of quantitative extension, dose

---

12 BHATTACHARYYA, D.: Some general characteristics of special libraries in science and technology in the U.K. In: Journal of documentation 28(3): 215, 1972.

13 LANCASTER, F. Wilfrid: Information retrieval systems: characteristics, testing and evaluation. New York, Wiley, (1968). p. 51.

14 LEAKE, Chauncey, D.: op.cit. p. 4.

related, of the ordinary actions of the drugs". [15]

It is difficult to draw a distinction between pharmacology and toxicology as all drugs produce side effects i.e. "anticipated effects on target organs other than those primarily intended". [16] When a drug produces abnormal side effects in an individual at a specific dose level, this abnormal side effect is termed an Adverse Drug Reaction.

3.21 User demand is the total of user needs.

3.22 User needs are those information needs of individuals that should be satisfied by certain specific information facilities.

3.23 User satisfaction is attained if the individual user of an information system obtains relevant usable information rapidly from a information system in response to his stated information need.

---

15 Loc.cit.

16 JACOBS, H.D.: Clinical evaluation of side effects of drugs. In: South African medical Journal 47(30): 1365, 1973.

CHAPTER ONE

THE DEVELOPMENT OF "SPECIAL" LIBRARY SYSTEMS

1.1 SCIENTIFIC COMMUNICATION

Communication in science has become complicated. The material to be communicated has increased tremendously both in depth and in breadth with the growth of science. Initially scientific communication was by word of mouth and by letter, both of which are today aided by technology and are consequently accelerated.

Scientific conferences, private meetings and discussions foster the communication of scientific information, as in the past, and scientific findings frequently make their way into print - the main medium for scientific communication for over three hundred years. Organizers and users of scientific information have to contend firstly with preliminary findings. Secondly, full publication may then take place in papers to a wide range of journals. Finally, there are secondary publications in abstract journals, reviews and journals announcing progress or advances in a particular subject.

Those who would safeguard the present flood of scientific information and explore means of improving scientific communication need to appreciate the extent of the volume of growth of this flood of information. Three groups have a vested interest in scientific publications. They are the producer i.e. the generator; the user; and the information scientist, or librarian. The producer of a 'communication' is often unaware of the existence of the information scientist/librarian and is

intent merely upon communicating with his fellow scientists. On the other hand, the user is often sadly unaware of sources for the information which he needs, or of the aid which the information librarian is able to provide.

A scientist is frequently unable to keep pace with anything more than a small fraction of the literature in his subject. The problem of keeping abreast of advancing knowledge in a subject is linked to two other problems - the retrieval of information already existing, and the transmission of relevant information from one science to another.

As mentioned above, the printed word is the main medium of communication in science and the scientific journals are the most important vehicles for this transmission of knowledge.

The most rapid, but restricted, means of scientific communication is direct verbal communication. Geographical restrictions in this information flow are, however, frequently overcome by the presentation of research results at international conferences, symposia or meetings. The printed word, as the most important medium of communication, makes its contribution to these scientific meetings in the publication of conference proceedings or abstracts which may become documented and speedily available through abstract indexing services to fellow researchers.

Thus it appears that often such communication is followed by the documentary records of the communicated knowledge but these records may be widely dispersed in many separately published documents, the retrieval of which requires many tools to meet the prospective user's

requirements.

The role of the telephone and telex in the communication of knowledge must not be overlooked, together with more recent technological developments such as magnetic and videotape recording. It is clear, however, that the tape recorder is the medium most frequently used to speed up the publication of conference proceedings. The use of this device in the editing of proceedings prior to publication is invaluable. The emergence of the computer is one of the most significant developments of all. Its importance is discussed on p.6-8 and in Chapers 6 and 7.

1.2 KNOWLEDGE, PEOPLE AND TECHNOLOGY

1.2.1 SCIENTIFIC LITERATURE

Scientific literature should be treated as a social phenomenon, for it is not only the product of knowledge and technology but, in addition, it is the product of both of the former and people. As the population explodes, the number of people attempting to communicate the scientific research increases and therefore the volume of scientific literature increases. This scientific literature is often managed by information scientists who need to anticipate the growth, demand and obsolescence of the literature.

A distinction needs to be drawn between knowledge of a subject and knowledge of its literature. The physicist, however well-read in physics he may be, does not necessarily know which indexes to published papers, theses or abstracts to consult in order to keep up with new developments in his field. He might scan or even read one or more recognized journals in his field but he is, nevertheless, frequently dependent upon others to provide access to the ever increasing mass of

published information in his subject field.

Libraries', librarians' and information scientists' former role was to collect the body of published material and store it for future use. Today it is no longer possible to collect indiscriminately and await possible use. It is neither economically nor practically feasible, because of the explosion of documentary information.

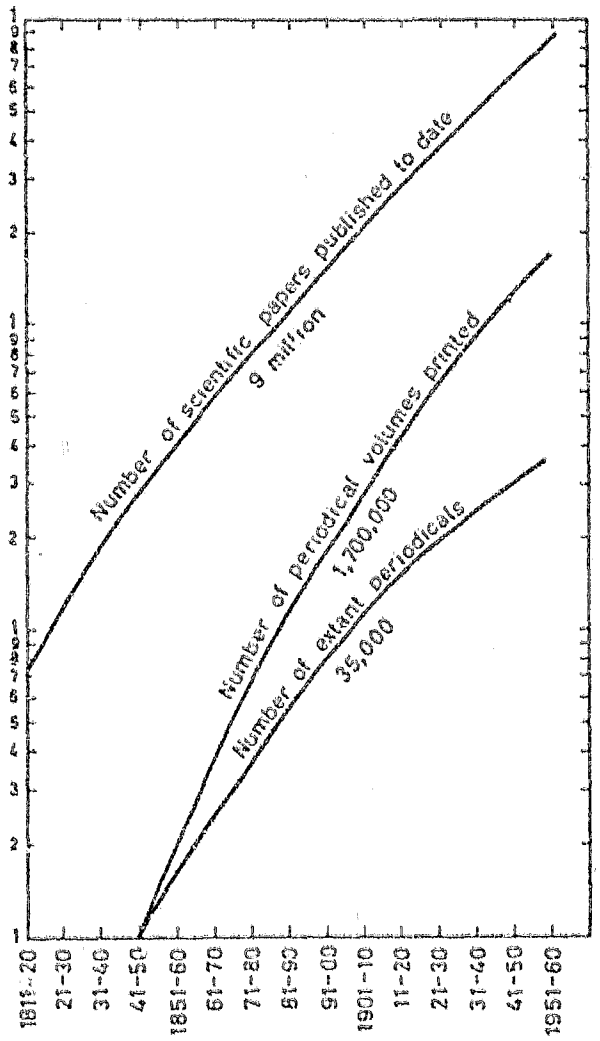


Figure I.1 Scientific Publication 1811 - 1960

One estimate of the growth of the mass of documentary information by B.C. Vickery is presented in Figure I,1 in which he shows that the total volume of documents doubles every twenty five years or less. [1.1]

The so-called "paper explosion" places a tremendous strain upon the often limited resources of people, time and money utilized to organize the mass of information. A consequence has been the development of the more sophisticated librarian or information scientist who evaluates the worth of much of today's research publication and finds the means of tailoring the library collection to meet the needs of its users and potential users. These information needs are not only current needs but also future or potential needs.

Currently produced literature is becoming an even greater problem in any subject speciality because of the increasing volume of journals published in different languages, the overlapping of fields of study or disciplines, the limited distribution of some reports, as well as the publication of papers in the ever widening range of journals. Whereas, in the past, a scientist would cover the handful of leading journals in his field, this has now become impossible and he needs to rely on current awareness services such as those provided by abstracting journals e.g. Chemical Abstracts, Pestdoc, Vetdoc, Ringdoc.

1.2 SPECIALIZATION OF LIBRARIES

The economic factor alone, makes it difficult, if not impossible,

---

1.1 VICKERY, B.C.: Techniques of information retrieval. London, Butterworths, (©1970). p.4.

to hold all the publications which may be of use to the library's users. This presupposes that each library will need to draw on the resources of others to satisfy user needs and that inter-library loan services will be utilised. The exclusion of a publication from the library stock is not always a measure of its value, however, as it may be better in another library's holdings - often because of subject content and frequency of use. A consequence of this is the development of highly specialized subject collections and the organized elimination of duplication within geographical areas or countries, together with the optimal use of the available resources.

1.2.3 SPECIAL LIBRARIES

King and Palmour [1.2] have reported that United States industrial chemists spend almost twelve hours per week in literature searching and current awareness reading. This indicates that assistance in speeding up this process would result in noticeable time and manpower savings. In an industrial environment where research and development are geared to producing profitable results, these savings become more important and valuable.

Computerization of information services in industry may benefit scientific workers b[illegible]creasing access to available resources and by decreasing the co[illegible] per unit of information stored. The expense of computerizing has, of course, to be justified. Computerization would

---

1.2 KING, Donald W. and PALMOUR, Vernon E.
In: Fenichel, C. ed.: Changing patterns in information retrieval. Washington, American society for information science, 1974. p.7 - 33.

not be feasible if the operational cost of running a search reached a prohibitive level. In a commercial or industrial concern the costs involved in not providing an information service must be weighed against the costs of manual versus a computerized information storage and retrieval system. (see Chapter 7). Even without such mechanical aids, however, industry has discovered that the absence or unavailability of necessary information at a particular moment can be costly. As a result, many industrial concerns are setting up their own libraries in order to prevent such occurences, as well as to save the duplication of work by centralizing the information store with consequent savings in cost, time and manpower.

Bhattacharyya has stated "special libraries... play a vital role in efficient information transfer and make their distinct contribution, in ways not yet fully understood, to creativity in basic science and innovation in technology. Indexing, or information storage and retrieval, plays an essential role in this transfer process". [1.3]

Special libraries are usually concerned with one subject or group of subjects and may have small holdings of material in borderline subjects. They may be libraries which hold collections predominantly consisting of one form of material e.g. maps or music. In industry, the special library exists to serve the needs of the staff of the organization, no matter whether they are production, marketing, research or management orientated. In an industrial special library much of the information store may be confidential and for internal use only.

1.3 BHATTACHARYYA, K.: op.cit. p. 214.

The essential characteristics of the special library are the types of material stored and the techniques used in order to retrieve information from the document store. Currency of information is of tremendous importance and for this reason a large proportion of the special library's budget is devoted to the acquisition of serial publications, including abstracting services. Careful selection and control of periodicals in practised and useful periodical articles are usually indicated by means of analytical index entries.

In the special library, the emphasis is usually placed upon creting a demand for information from the library, and the professional staff may compile subject interest profiles for the library's users and alert them to incoming potentially interesting items of information i.e., institute a selective dissemination of information (SDI) service. Other current awareness services such as a frequent circular or newsletter of accessions lists of new books, periodicals and reports may be instituted. The quantity of current material received enforces the necessity for this type of service in the special library. The special library staff have detailed knowledge of staff activities as well as of the organizational goals of the organization which the library serves and, as a result, user interest profiles are comprehensive in coverage and may be regularly updated to keep pace with changing emphases within the organization.

The problem of coping with the mass of material in even a restricted subject field is often overwhelming and leads to the fact that it is frequently the special library which innovates in information storage, handling and retrieval. New techniques are tried and mechanical aids adopted in order to contain costs and control the volume of information handled.

The extensive subject records necessary in the special library require careful subject indexing for retrieval purposes. Speeding up of this process by mechanical means may free the information subject specialist to assist the user in meeting his information needs: Classification and indexing are important in the special library and need to be detailed in order to provide in-depth analysis of pertinent subject fields. A number of access points to the specialised subject material need to be provided in a library with a subject approach.

Because of the highly specialised subject needs of the clientele of the special library, co-operation with other libraries assumes an important rôle. Even in a limited subject area, it is impossible to purchase all the current material being published and to be self-sufficient and able to meet all the user's information requirements. Co-operation in the pharmaceutical industry information field is discussed in Chapter 7.

The special library must save the user time and this may be accomplished in a number of ways. Firstly, user time spent in reading for items of general interest and in order to keep up to date in a specific field may be decreased by use of the library's current awareness services.

Secondly, the growing use of subject specialists in the special library who are equipped by their training to sift through the information and select relevant information for submission to the user saves the user much time usually spent in unproductive browsing. Not all browsing is unproductive, however, and browsing may be rewarding or valuable if the user is seeking inspiration.

Thirdly, rctrospective information searches may not be left to the user himself to do, but may be undertaken by the library staff. Lancaster calls this the "delegated search" i.e., a search "delegate... by the person having the information need, to a second individual, usually a librarian or information specialist." [1.4] These delegated searches tend to be "iterative" [1.5], in other words, the information specialist conducts a search on behalf of the user and interacts with the user during the search procedure in order to verify the relevance of retrieved documents, and either may modify the search strategy according to this interaction.

Fourthly, periodicals are circulated to the user. This function of the special libra.y usually forms part of the current awareness service, as items of particular interest to individuals may be selected and page locations indicated on the periodical circulation lists. Various bibliographies of current publications in a specific subject field may be drawn up in anticipation of the user's future needs. These services are not always in response to a stated interest and may be prompted by the information worker's knowledge of the environment in which the library service operates. This function of the special library may not operate on a large scale and may be restricted to certain key users of the service.

Most special libraries institute current awareness services. These may vary widely in format, coverage and currency, but form an integral part of the services offered and assist in the user time-sparing activities of the library service.

1.4 LANCASTER, F. Wilfred: op.cit. p. 181.

1.5 Ibid. p. 182.

This is not to negate the useful browsing potential of any collection, however, but to draw to the attention of the individual those seemingly useful bibliographical references, titles, abstract or books to which he might attach a value judgement [1.5] of adequacy or usefulness.

Organizers of current awareness services must take cognisance of the fact that an individual's preferences, expectations, purposes and needs change.

1.5 DEWEY, J.: The field of value.
In: LEPLEY, R. ed.: Value, a cooperative inquiry. New York, Columbia univ. press, 1949. Discussed in SWANSON, R.W.: Performing evaluation studies in information science. See bibliography.

CHAPTER TWO

AN OUTLINE OF THE LIBRARY SYSTEM WITHIN CIBA-GEIGY (PTY) LTD., SOUTH AFRICA

2.1 THE BACKGROUND AND STRUCTURE OF CIBA-GEIGY LTD.

2.1.1 INTRODUCTION

Ciba-Geigy is a multinational organization which originates, manufactures and markets organic chemical specialities. The group is made up of affiliates in fifty countries and has 76,000 employees around the world.

The parent company is Ciba-Geigy Ltd. of Basle, Switzerland, formed in October 1970 by the merger of two long established companies Ciba A.G. and J.R. Geigy A.G.

Because of Switzerland's lack of natural resources, the company specializes in original high-quality chemicals and this accounts for its emphasis on research. The activities of the company cover pharmaceuticals, dyestuffs, agrochemicals, plastics and additives, dyestuffs and pigments, consumer products and photographic materials. Each of the first five listed spheres of interest is concerned with its own research and development, production and marketing, with most of the research and development being undertaken in Basle, Switzerland.

Co-ordination between central management and the fifty international affiliates is achieved through six regional service departments covering Europe, Latin America, the Near East and Africa, Eastern

Europe and the Indian Subcontinent , the Far East and Australia. [2.1] Figure II,1 the Ciba-Geigy Organization Chart, illustrates seven regional services, six divisions and six central functions.

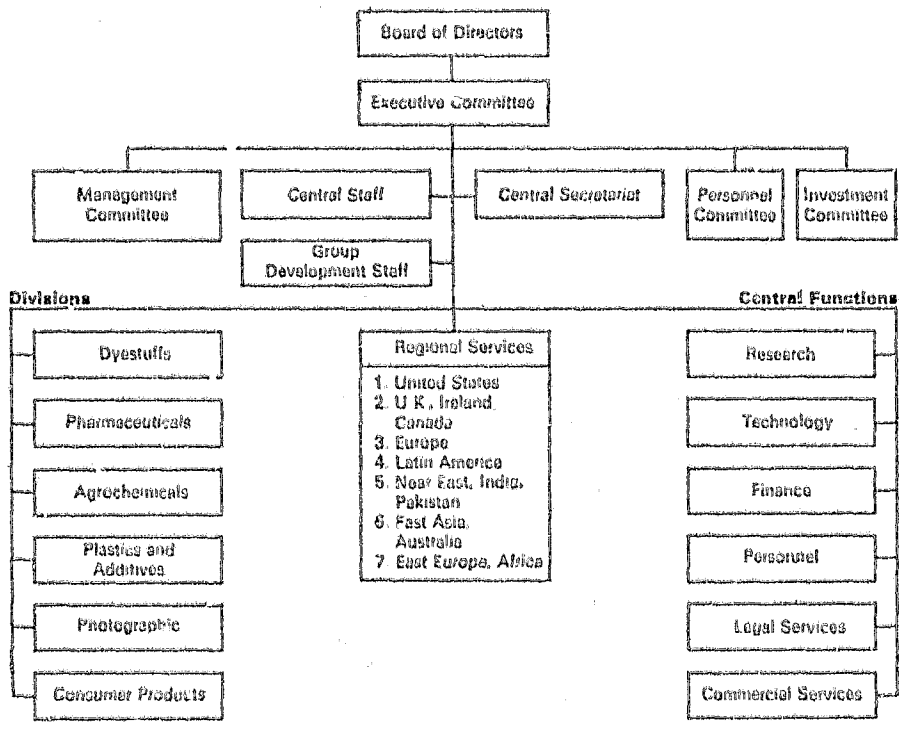


Figure II,1 Organization Chart CIBA-Geigy Limited

The Ciba-Geigy group is considered an economic entity but, for administrative reasons, it is divided into parent company and group companies, and subdivided into divisions. Matters of divisional interest are dealt with within the division concerned. This is of particular importance in the dissemination of information between the parent company and group companies, and affects the organization

---

2.1 OATES, David: Ciba-Geigy reassesses its global future. In: International management 28(3): 17, 1973.

of the South African group company's library system. This will be dealt with later in Chapter 2.3 The development and functions of the library system.

2.1.2 CIBA A.G.

In 1859 a Swiss industrialist, Alexander Clavel, started manufacturing a coal-tar dye and in 1884 the enterprise thus started was incorporated as "The Society of Chemical Industry in Basle", later abbreviated to CIBA, which was adopted as the official name of the company. In 1889 the Society exhibited its first pharmaceuticals at the World Exhibition in Paris. Further rapid growth ensued during the following three decades as a result of development in the dyes and drug fields.

CIBA entered the plastics field in 1936, while the marketing of a dentrifice in 1952 heralded its entry into the consumer goods field. The company's position was further strengthened in the pharmaceutical field by the isolation and partial synthesis of hormones in the early 1930's. In the 1940's CIBA marketed the epoxy resins.

In 1965 three new divisions, Plastics, Agrochemicals and Photochemicals were created out of the former Technical Synthetics section. Economic and technological growth was accompanied by international diversification, with subsidiaries established world-wide.

The Ciba Foundation for the Promotion of International Co-operation in Medical and Chemical Research was founded and endowed by the organization, and its symposia "have become essential references

in the libraries of the world. [2.2] Titles in the series include:-

Bilharziasis
Carcinogenesis
Cardiomyopathies
Caries-resistant teeth
Cellular aspects of immunity
Civilization and science: in conflict or collaboration?
Congenital malformations
Development of the lung
Disorders of language
The family and its future
Foetal autonomy
Lipids, malnutrition and the developing brain
Physiology, emotion and psychosomatic illness
Intrauterine infections
Atherogenesis: initiating factors
Law and ethics of A.I.D. and embryo transfer
The poisoned patient: the role of the laboratory
Size at birth
Health and industrial growth
Parent-infant interaction
Research and medical practice: their interaction
Breastfeeding and the mother
The fetus and birth
Health and disease in tribal society
Respiratory tract mucus

2.1.3 J.R. GEIGY S.A.

In 1758 Johan Rudolf Geigy established a dyes and chemicals company. He was a trader in 'colonial' goods - coffee (then regarded

---

2.2 CIBA
Basle, Ciba, 1967. p. 14.

as a drug), pepper, cocoa, tea, spices, indigo, the tanning agent suniac, and opium. His grandson, Carl Geigy, concentrated on dyes, especially indigo and madder. In 1863, Carl's son, J.R. Geigy-Merian employed chemists for the production of synthetic dyes for the first time.

Geigy progressively enlarged its market to include India, China, Japan, the Phillipines, Java, the U.S.A., France and Russia. In 1938 a pharmaceutical division was added and the agricultural chemical field was entered in 1948 when Dr. P. Müller won the Nobel prize for Physiology and Medicine for his discovery of the insecticidal properties of DDT.

Geigy's most important affiliates were in the U.S.A. and Britain. As with Ciba, five regions were established to co-ordinate group activities, with six functions serving as support activities. (see Figure II,2.)

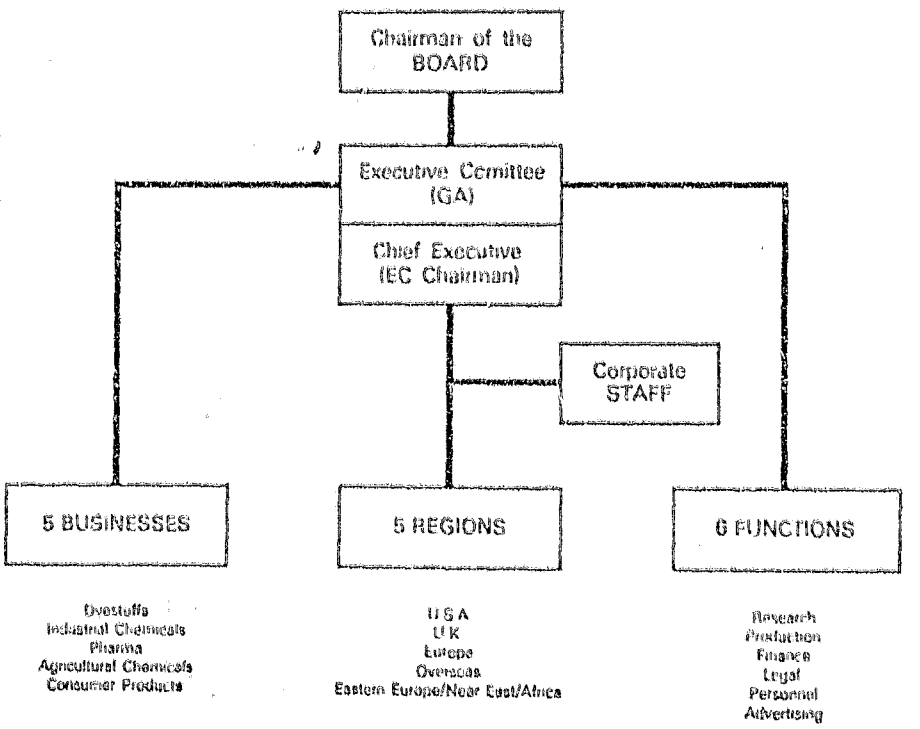


Figure II,2 Geigy, Basle Organization Chart

2.1.4 RESEARCH DEVELOPMENTS

The importance of research by private enterprise for progress in the field of pharmaceuticals has been shown by American investigations.[2.3] The number of new drugs introduced has, however, shown an appreciable decline. The trend is due mainly to the following causes: ... "the relatively high level of development that the pharmaceutical industry has attained in many fields renders it increasingly difficult to find new, improved drugs for conventional indications. Whereas only a few years ago it could be regarded as reasonably certain that one out of a thousand new chemical compounds investigated would be worth while introducing into therapy, this ratio ... has now risen to 1:4,000. On the other hand, the demands of new drugs in clinical trials have become much more exacting. Prior to introduction, a new drug nowadays sometimes has to be tested in 10,000 or more patients because reasonably accurate statistics on the incidence and severity of side effects can only be [illegible]red if very large groups of patients are treated. The steady in[illegible]se in experience and in medical and biological knowledge i[illegible] paralleled by an increase in the number and variety of the tests [illegible] which a responsible firm has to submit a new drug prior to introduction."

"Incidentally, research into a drug does not cease with the latter's commercial introduction but continues for many years afterwards.

2.3 UNITED STATES. Subcommittee on antitrust and monopoly. Study of administered prices in the drug industry. Washington, 1961, p. 115 - 118.

In: Problems facing the research-based pharmaceutical industry. Biel, Ciba, Geigy, Roche, Sandoz, Wander, 1966. p. 24.

Observations on effects or side effects which do not appear until the drug has been widely used are reported back to the research department by clinicians and general practitioners and new trials are often required to analyse and clarify these observations. In some cases, new and entirely unexpected indications for the use of the drug are found years after its introduction." [2.4]

An example of a new indication for an old drug is seen in the case of sulfinpyrazone which was first marketed in South Africa in 1961, and in mid 1975, the new therapeutic indication "clinical states in which abnormal platelet aggregation and survival are causative or associated factors" [2.5] was added.

A clinical trial has been defined by Dollery [2.6] as "a scientific experiment conducted in man with the aim of improving the treatment of the sick."

Clinical trials are generally divided into four phases. Phase I studies test the initial administration of the drug to man in small groups of volunteers or patients. The pharmacokinetics of the drug are measured. Phase II studies cover the initial trials on a limited number of patients for specific disease control. Phase III studies provide the assessment of the drug's safety and effectiveness in groups of subjects involving a given disease or condition. Phase IV studies are studies which may be used to comply with government regulations

---

2.4 Loc.cit.

2.5 MIMS desk reference v.12, 1976 - 1977.
Pretoria, MIMS, 1976. p. 230.

2.6 DOLLERY, C.T. and BENNETT, P.N.: Clinical trials.
In: British Journal of clinical pharmacology 2(6):
479 - 480, 1975.

involving the introduction of new drugs into the market and may emphasize drug efficacy or adverse drug reaction monitoring.

This research may be organized, in the form of clinical trials, which are scientifically monitored to evaluate drugs in the treatment of disease, or it may be in the form of personal observation by specialists in various fields of medicine. The latter method has, however, fallen into disrepute because of its subjectivity and has virtually been superseded by the clinical trial.

"The rational use of a drug ... by the health professions, in agriculture or in other ways is dependent on the co-ordinated research endeavors of many kinds of scientists, chiefly chemists, pharmacologists and clinicians. ... These also appropriately are the ones on whom the general public must rely for satisfactory advice [information] on the rational use of a drug and for effective control of its possible abuse." 2.7

The development of new drugs involves the storage of information on the following aspects of the drug, its chemistry, manufacture, including quality control, (Step 1). (see Figure II.3 Drug Development and Stages in the Recording of Information about the Drug. p. 20). The second step is the recording of information on the pharmacology and pharmacokinetics of the drug - its absorption, metabolism and excretion. Thirdly, and linked with the pharmacology of the drug, the recording of information on its toxicity. The drug may then be released for clinical evaluation (Step 4) in either hospitalized or ambulatory patients (Step 5). Information (pharmacology or toxicology) gleaned from the evaluation in either group of patients, is

---

2.7 LEAKE, Chauncey: op.cit. p.9.

either passed back to the clinical trialist who advances the information to the manufacturer (pharmaceutical company), or is gathered by the manufacturer directly from clinical trial record forms of each patient entered into the clinical trial.

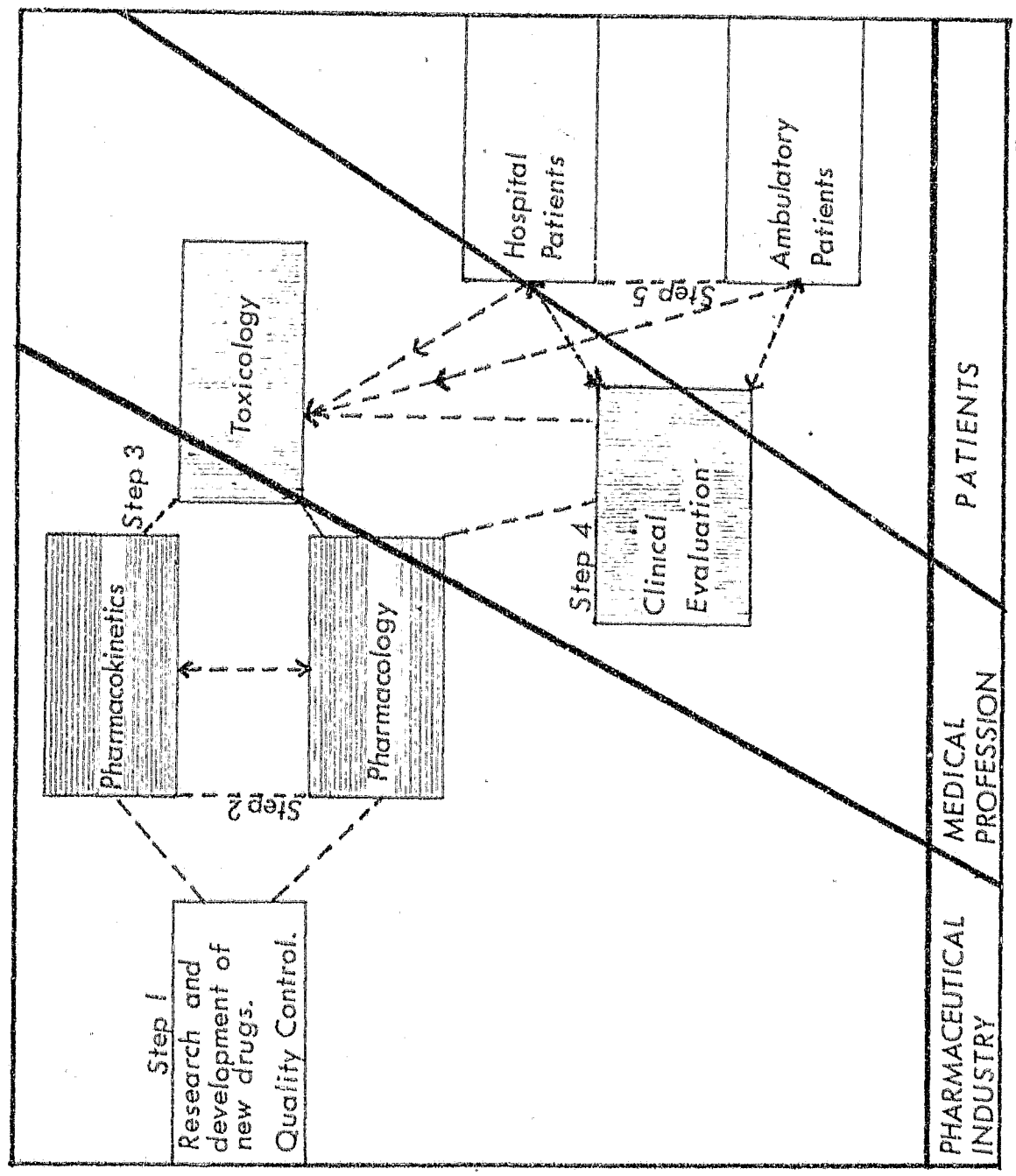


Figure II,3 Drug Development and Stages in the Recording of Information about the Drug.

Professor H.W. Snyman, former Chairman of the South African Medicines Control Council, in an address has ranked new medicines into five groups "in terms of their total therapeutic value - i.e. adequate efficacy and adequate safety" [2.8]. These groups are:

"(a) The fundamental new medicine of major clinical significance;

(b) An important new medicine, offering substantial advantages for a majority of patients;

(c) A useful new medicine, offering substantial advantages for a minority of patients;

(d) A new medicine, offering only marginal advantages over previously available preparations;

(e) A new medicine, offering little or no advantage over previously available preparations." [2.9]

The number of available drugs is ever-increasing and places a tremendous strain on the doctor who has to manage the mass of information flooding his surgery daily on new and old drugs. This information often mirrors the expansion of knowledge about the pharmacology of drugs or the physiological basis of therapeutics. The application of scientific methods in the doctor's choice of a drug to be administered to a patient requires knowledge about the iatrogenic possibilities of the drug. The doctor must be aware of contra-indications to the use of the drug, about drug interactions with other drugs, and with foodstuffs e.g. cheese and monoamine oxidase inhibitors.

2.8 ORKIN, Carole J.: Growing problem of price and prescription medicines.
In: Medical chronicle Nov. 1976. p.6.

2.9 Loc.cit.

"The physician is under an obligation to prescribe on a rational - as opposed to an irrational or intuitive - basis not only for ethical and scientific reasons but also out of a sense of social and socio-economic responsibility towards his patient or the patients' insurer." 2.10

The role of the pharmaceutical industry is, therefore, to place this information at the disposal of the medical profession. It must not be forgotten that the library is frequently called upon by the sales department to provide information of a promotional nature on the company's drugs, often used in advertising and the promotional mailing. The first phase in the composition of promotional claims by the drug manufacturer begins in the information centre, where published claims as to drug efficacy and safety are sought. (see Figure II,4 p. 23).

Schwartzman states that "since drugs have many quality dimensions, the quantity of useful information about each of them is large. A drug is a complex product, the effects of which may be numerous, and these effects vary among patients. To use a drug well, doctors must be able to anticipate these effects accurately. Hence the demand for information concerning the quality of drugs by doctors is much greater than the demand for information by consumers generally concerning the quality of other products. Doctors' prescriptions for a drug will therefore tend to increase with the quantity of information which they have about the drug for any given level of quality and price, and manufacturers will have an incentive to supply information." 2.11

2.10 CERLETTI, A.: Drugs, the individual and society. Basle, Pharma information, 1970, p.8.

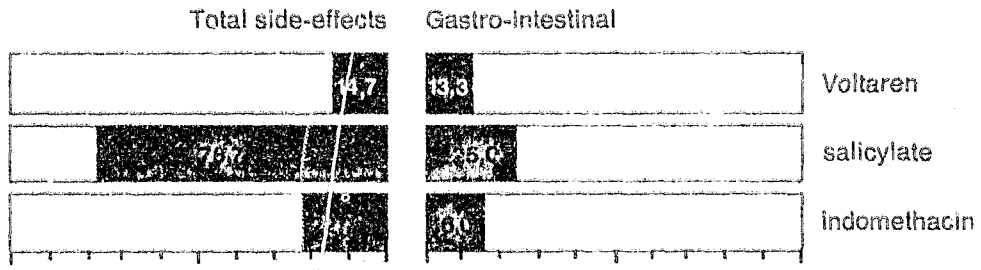
2.11 SCHWARTZMAN, David: Innovation in the pharmaceutical industry.
Baltimore, Johns Hopkins univ. press, (c1976) p. 185.

®Voltaren Geigy

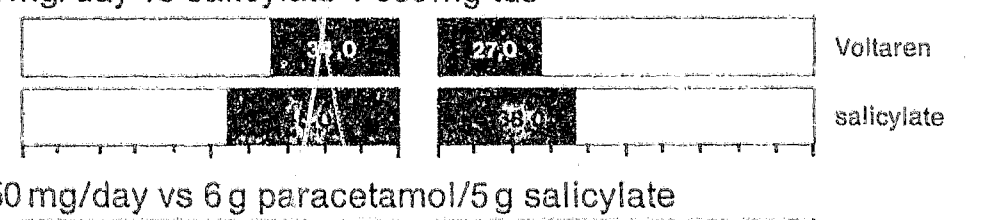
provides better patient tolerance

"However, VOLTAREN produced significantly fewer side-effects than were encountered during either salicylate or indomethacin administration"[1]

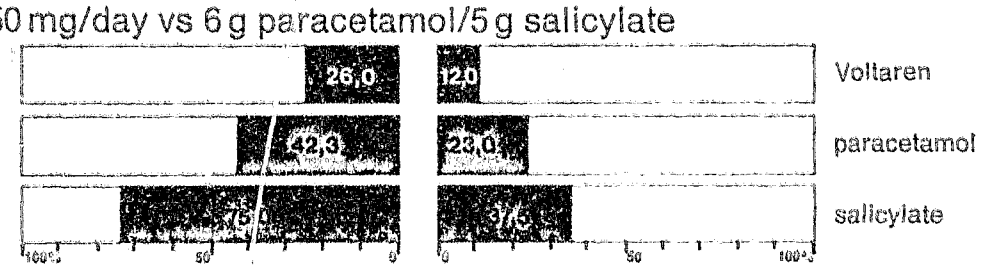
Voltaren 75mg/day vs salicylate 4,5g/day and indomethacin 75mg/day



Voltaren 75mg/day vs salicylate 1 000mg tds



Voltaren 150 mg/day vs 6 g paracetamol/5 g salicylate



These South African clinical findings have conclusively demonstrated that:

In Rheumatoid Arthritis[1,2]:

- VOLTAREN was statistically superior or equal to salicylate, indomethacin, paracetamol in effectively relieving early morning stiffness.
- VOLTAREN was preferred when compared to salicylate in treating Rheumatoid Arthritis.

In Osteo-arthritis[3]:

- VOLTAREN was superior to salicylate in treating this condition.
- VOLTAREN produces less side-effects than salicylate, paracetamol or indomethacin.
- VOLTAREN is strictly controlled by prescription – thus enabling the doctor to supervise therapy.
- VOLTAREN is a research product of Geigy – leaders in arthritic therapy.

Usual Dosage: 25 mg t.i.d.

References:
1. Meyers, O. L., et al.: *S. Afr. Med J.*, 48(48): 2013-2017, 1974.
2. Solomon, L. and Abrams, G.: *S. Afr. Med J.*, 48(22): 949-952, 1974.
3. Joubert, P. H., et al.: *S. Afr. Med. J.*, 48(47): 1973-1978, 1974.

Figure II,4 Promotional Mailing Showing Efficacy and Tolerability Data Obtained from the Library

Schwartzman [2.12] points out that critics of the pharmaceutical industry urge doctors to balance the conflicting claims of manufacturers of competitive drugs by using alternative sources of information such as articles in medical journals and expert opinions of colleagues. He provides some interesting statistics on medical journal articles. "Since 1970 an average of 1700 articles on any one of the twenty-five leading drugs has been published per year in 325 journals of medicine, pharmacology, chemistry and pharmacy. This astonishingly large number underestimates the size of the world literature on each of the major drugs, since there are other journals not even counted in the estimate. [2.13]

It is obvious, therefore, that the clinical innovators who expend large sums on original research will use scientific publications to sell their preparations. "It has been pointed out that developing a new drug may cost $ 55 million, and that only 25 laboratories around the world can carry such a financial burden, so that eventually the larger pharmaceutical firms are likely to be the only innovators." [2.14]

Unsolicited information on drugs is thus diffused to medical practitioners by means of the promotional mailing. This drug advertising practice is common in South Africa, although much of the information on drugs received by mail is not presented in such a way as to be retained by medical practitioners, it forms one of the

---

2.12 Ibid. p. 186.

2.13 Loc.cit.

2.14 PESSIMISM about drug research.
In: South African medical Journal 53(17): 652, 1978.

principle means by which the pharmaceutical industry makes its new drugs known to the medical and pharmacy professions. (see Figure II,4 p.23). This new product, Voltaren, an antirheumatic drug, underwent Phase III clinical trials in South Africa. The resultant publications (References 1 - 3) were utilized in this promotional mailing to provide information on the comparative efficacy and tolerability of the preparation. Incidentally, these papers were brought to the attention of sales and marketing staff by means of the library's current awareness service. The subsequent use to which the information was put is seen in Figure II,4 p. 23.

In addition, the pharmaceutical industry often supports continuing medical education by means of informative monographs on advances in therapeutics, and, more recently, by means of audio-visual programmes. The preparation for local audio-visual programmes is, to some extent, dependent upon the published results of clinical trials from the world literature, and it is for this reason that the company's information system plays a role in the collection of new results from clinical research and their dissemination to a wider audience.

Continuing medical education is supported by the company, by the distribution of publications such as the 'Ciba Clinical Symposia', the 'Documenta Geigy Folia Traumatologica' and 'Acta Rheumatologica', as well as the 'Ciba Foundation Symposia' (selected titles of which are listed on p. 15). In addition, Ciba-Geigy publishes or distributes the 'Documenta Geigy Scientific Tables' and the 'Ciba Collection of Medical Illustrations' by Frank Netter, both of which are widely used and form standard reference works in many medical and scientific libraries and in private collections.

2.2 THE BACKGROUND AND STRUCTURE OF CIBA-GEIGY (PTY) LTD., SOUTH AFRICA

In 1958 Ciba (Pty) Limited was established in South Africa and was located in Johannesburg. Formerly, it had existed in Cape Town as an agency for Ciba Limited of Basle, Switzerland. It had two divisions, pharmaceuticals and dyestuffs and, by 1966, agricultural chemicals were also marketed.

Geigy operated in South Africa as Pharmakers from 1950 to 1962 and marketed Geigy pharmaceuticals. Geigy (Pty) Limited was established in Johannesburg in 1958 and, in 1968, Pharmakers merged with Geigy. By 1968, agricultural chemicals, dyestuffs, and plastics and additives were marketed in addition to pharmaceuticals.

Following the merger of the parent companies in Basle, Switzerland, Ciba and Geigy merged in December 1970, although the formal merger date was 1st January 1971.

In 1971, the company structure was as follows:-

Four divisions :- pharmaceuticals; dyestuffs; pigments and additives; and agricultural chemicals.

Two functions :- personnel; and finance and administration.

A further division was added in 1972, viz. the photographic division. A consumer products project was embarked upon in 1972.

Each company division is led by a division head who is responsible to the general manager. Divisions are, in turn, subdivided into departments controlled by departmental heads.

2.3 THE DEVELOPMENT AND FUNCTIONS OF THE LIBRARY SYSTEM

2.3.1 INTRODUCTION

"The purpose of a company library is to provide information that supports the organization's work and furthers its goals. In order to do this, the library must have a close identification with, and knowledge of, the organization - its structure, personnel, products, policies, plans and the political interaction of people and departments ... Dependence upon a 'trickle down' process for keeping the library informed about company plans and policies will eventually result in an information service which is irrelevant to the company's needs." 2.15

2.3.2 THE DEVELOPMENT OF THE LIBRARY

Before the merger of Geigy and Ciba at the end of 1970, the collections of material had been superficially organised by untrained staff in the South African Ciba embryo 'library'. The emphasis was placed on the book collection consisting predominantly of Ciba Foundation Symposia and Ciba Foundation Study Group publications on various medical topics. Certain Ciba product literature, usually in the form of reprints, was grouped together under the various product trade names. The strength of the collection lay in its microfilm holdings supplied by Ciba headquarters in Basle, Switzerland. Published papers on Ciba products were microfilmed. The film was cut into strips and each filmed document was inserted into a specially designed pocket on a 21cm X 15cm card on which was printed the complete bibliographical

---

2.15 MCHELMAN, Shirley: Libraries are businesses, too. In: Special libraries 65(10 - 11): 409, 1974.

citation, an English abstract of the contents of the document, as well as all the indexing terms assigned to the document. The indexing was done manually in company headquarters in Switzerland and the chosen terms fed into a computer which stored the material and sorted it into author, product-subject, side effects, and abstracts/or summary sequences. Annual computer-printed indexes were then issued for each sequence.

Geigy South Africa, in contrast, had a growing, organised pharmaceutical division library when the merger of the two companies took place. This library had a small book collection and drew to a large extent on the book stock of the local medical school library. Its product information was supplied on microfilm by its company headquarters, also in Basle. Small envelopes with author, title and source were provided and local staff were required to cut the microfilm into strips, inserting the strips into the correct envelopes. This procedure was time-consuming as the envelopes then were sorted alphabetically, by author, under each product trade name. There was no index to this material. Simultaneously, however, the material on microfilm was supplied in photocopy form, much of which was discarded as it was either of limited scientific or medical value - e.g. a medical practitioner's subjective opinion of the clinical efficacy of a drug - or there was a language barrier to its use in South Africa. Languages accepted were English, German, Dutch, Portugese, Spanish and French.

A number of medical and pharmaceutical journals were subscribed to, but prior to the organization of the library by a librarian, these journals had been destroyed after circulation so the only material available for retrospective searches was the microfilm of photocopied product literature.

After reorganization of the library, all journals were scanned and items of interest were photocopied for storage in indexed author, subject, and product sequences. The journals themselves were bound for use in retrospective searches.

In 1971, after the companies' merger, it became necessary to integrate these two collections of material. Ciba-Geigy headquarters in Basle adopted the Ciba product documentation system and supplied published papers on microfiche with an annual index. There was to be no cumulated index to this material. Locally, the Geigy subject list of indexing terms was used to integrate those documents stored in the company's pharmaceutical division library.

2.3.3 THE FUNCTIONS OF THE LIBRARY

The company library must collect, arrange and organise its material so as to render it quickly and conveniently available for inquirers. Its function is to liase between research on the one hand and practical business on the other hand. The emphasis in a company library falls more heavily on 'inquirers' than on 'readers' and the library staff are expected to provide direct answers to inquirers without the written word interposing necessarily. In the company library, the element of 'anticipation' is important, as business trends and interests should be anticipated and material accumulated in advance in order to ensure immediate availability at any future time. This, of course, may save the company financial expenditure and aid its profit-making objective.

The company library must, therefore, be geared to provide information rapidly often in considerable detail in order to solve a

multitude of diverse problems often related to the business operations of the company and sometimes of a highly technical nature.

There is a divergence of opinion as to whether a company library has a didactic function - to promote informal education amongst company employees. The manner in which a company library is integrated into company objectives is not always clear. In some companies company objectives may be ill-defined and this may pose problems for the interlinked library in defining its own objectives and serving or anticipating the company's information needs. The ultimate aim of any library is the service of the user and it is, therefore, important to consider how information is dealt with, what that information is, who it is intended for and to what extent the dissemination of this information assists in the spread of scientific and technical knowledge.

One of the greatest problems in serving the user is the problem of the user's requirements from the information service. Frequently the user knows what he wants, but not what he needs i.e., he has incompletely analysed his requirements. It requires competent inquiry staff to distinguish the information need underlying the user's request.

Before satisfactory answers can be given to the user's request, the 'input' - the search strategy - should be formulated according to a logical analysis of the user's requirements for that particular information need. The user's verbalization of his information need must be examined and interpreted in order to establish its essence, the searcher must then convert the terminology expressed by the user into that in use in the information system and adjust the search strategy according to the system terminology so as to increase the

possibility of the retrival of relevant information.

Ineffective communication at the search strategy formulation stage decreases the likelihood of retrieving relevant information, documents, or document surrogates. The onus is upon the information staff to facilitate effective communication in order to improve recall and precision in the retrival of material from the information system.

The library must plan for increases in user demand on the system which could be caused by increases in the number of staff employed by the company as well as by the company's increased activity in clinical research. User demand may also be increased by more demands per capita from existing direct users (staff members) or by an increased awareness of the services offered by the library among the potential user population, chiefly, the medical profession (see Chapter 7).

2.3.4. INFORMATION FLOW IN THE PHARMACEUTICAL INDUSTRY

As can be seen from the flow charts below (Figures II,5 and II,6 p.32-3) observations about a drug are continually noted and are communicated from marketer to researcher and back again. There is thus an ongoing diffusion of information by the pharmaceutical industy to the medical, para-medical and pharmacy professions and vice versa.

The need to supply information on (a) products and (b) research substances may be examined by undertaking a user survey of information requests and by analysis of the distribution of the information.

Analysis of the distribution of the information may indicate

areas within the company itself in which needs are not being met and will identify external groups which are receiving information from the company library. (see Chapters 3 and 7).

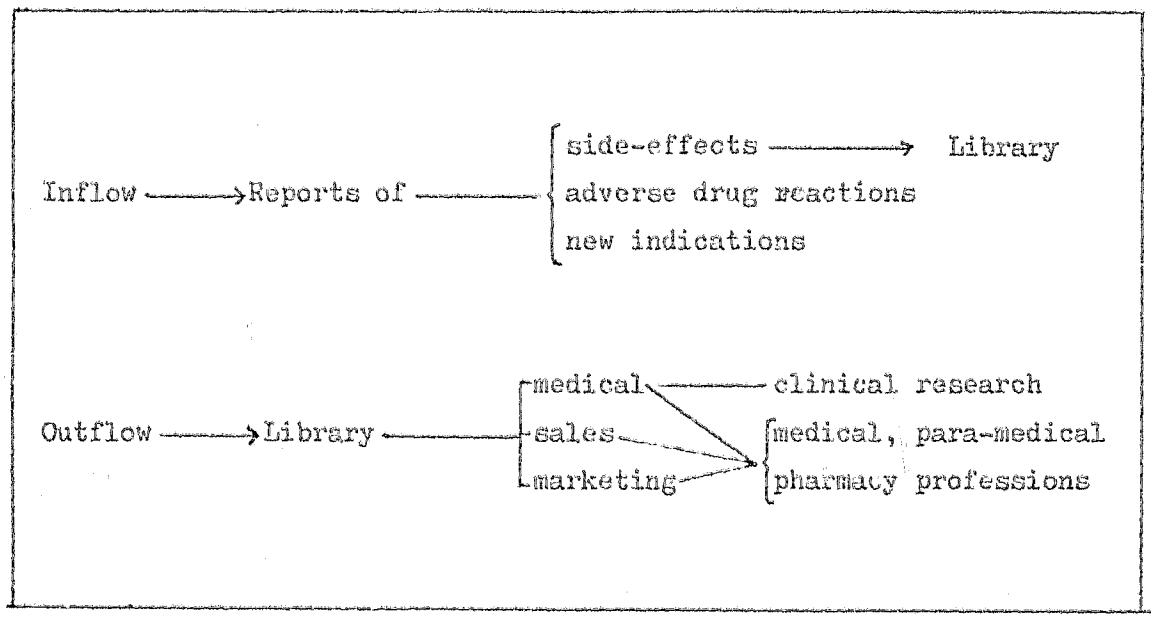


Figure II,5 Drug Information Flow

As discussed on pages 18 to 20, this inflow is frequently the feedback from clinical trials (usually Phase IV studies) in which drug safety, side effects of drugs, adverse drug reactions and new indications for a drug are recorded in the library from published sources or unpublished reports.

The information thus recorded, together with information from other sources, e.g. journals, abstracting journals and books, forms the information store which certain departments within the company draw upon in order to answer queries from outside the organization (chiefly, from the medical and pharmacy professions). (see Figure II,5) In addition, the company's medical department may seek information from the library in order to initiate clinical research or to brief clinical trialists on a product. (see Figure II,7 overleaf)

Information on pharmaceutical products was chanelled

(a) for internal use (see Figure II,5. p.32, Figure II,6 below)

(b) for outside diffusion (see Figure II,7 below)

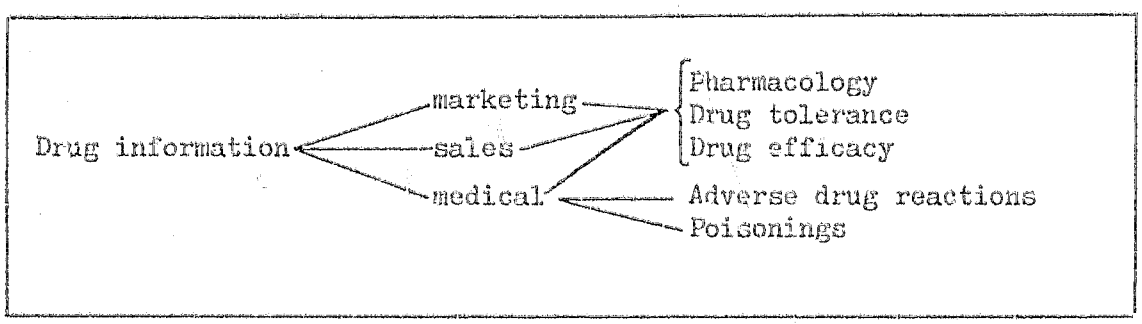


Figure II,6 Drug Information Flow Within the Organization

Information on research substances was diffused according to the flow chart in Figure II,6, and Figure II,7.

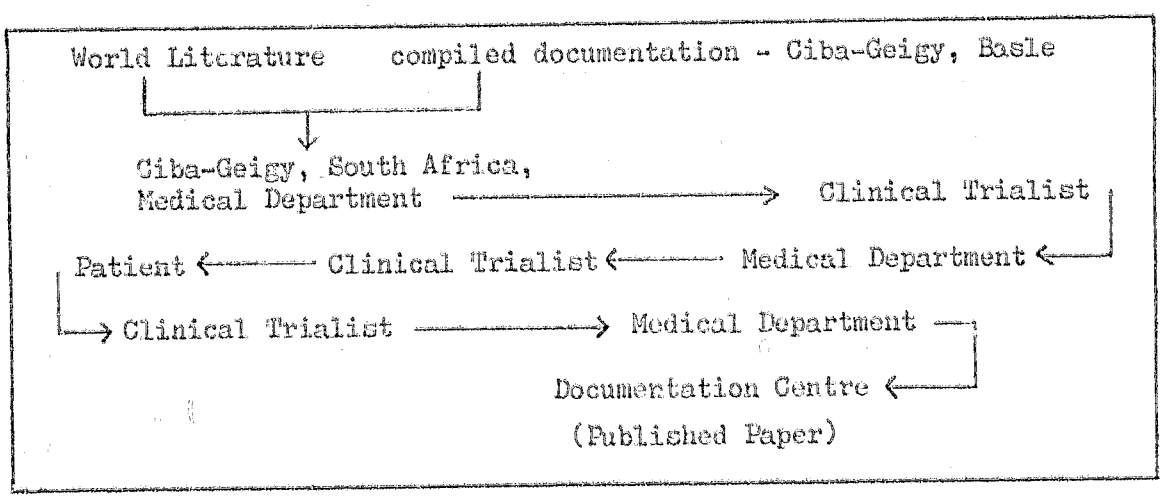


Figure II,7 Information Flow Prior to, During and Subsequent to Company Instituted Clinical Trials in South Africa.

An important objective in the pharmaceutical industry is to keep abreast of possible new indications for the use of drugs (see p. 18) and it is inevitable that the mass of information gleaned from medical practitioners and researchers should be organized by an information function or special library. In addition, each pharmaceutical concern has

found it essential to have immediate information about opposition drugs from other companies, either for their market research operations or for their own product management. In South Africa, not all pharmaceutical companies entrust the storage and retrival of this information to a company library and the information function may be divided amongst different functions.

It is essential that local experience with the use of the company's drugs be recorded and readily available should the need arise. An interesting example of this was the discovery during a clinical trial that the Indian population group metabolised a part··· r drug differently from both the White and Black groups. Dosages were noted and this information was recorded in the library system in its published form.

Outside inquiries from the medical profession are noted and consequently the input to the library system is influenced by the emphases thus detected. This flow of information to the medical profession thus exerts a formative effect on the information to be stored by the library, and illustrates how user needs influence library holdings and information storage. Users have different needs at different times and as these needs change so too do the library's holdings to accommodate the shifting emphases.

It is advantageous to the company's overseas affiliates and research associates to have all local serial publications and conference reports and proceedings scanned and any information in fields of interest immediately sent to a central Documentation Centre in Basle, Switzerland which then decides the merits of disseminating the information to group companies.

The advantage of a local library to the medical profession is that it provides more comprehensive information and/or bibliographical citations on pharmaceutical products, about which specific, but often fairly limited information is already known. (See Chapter 2.1.4 p. 17 - 25)

B.C. Vickery has succinctly summed up the reasons for the need for information retrieval.

> "Modern society incessantly produces and uses information. All technical activity - in science, industry, commerce or government - now takes place in such a complex environment that it mush be based on especially acquired information. At the same time, every act gives rise to information, and recorded knowledge grows apace. To find the information, one seeks within the huge mass now available becomes ever more difficult. If information is to be accessible it must be organized." 2.16

Organization of information - this is a factor which promotes both the effectiveness and efficiency of the dissemination of information and is as essential in the medical field as in any other field of research.

2.3.5. USER NEEDS

The scientist has a basic information need and that is to be kept informed of current advances in his field or fields of interest, in order, firstly, not to duplicate work which has already been done, and secondly, to enable him to build upon any new advances and achievements.

2.16 VICKERY, B.C.: op.cit. p.1.

"The year 1971 saw a continued growth of literature reporting information needs and uses in science and technology." 2.17

"Those who need scientific and technical information are scattered throughout every country in the world - in universities, research stations, factories, plantations, hospitals. To each user we must give an equal opportunity to identify the literature he needs and to obtain it without too much delay." 2.18

"The purpose of any retrieval system is, of course, to deliver output to users - ideally, just the information they need, in the form and at the time they need it. Before a system can be designed - before its input can be specified and its procedures decided - the required output must be determined. Essentially, this requires a study of user needs." 2.19

It is necessary to find out who these users are, in fact, also who the potential users might be. One way to do this is to undertake a user survey to determine within the institution the volume of demand from each type of job within the organisation and to evaluate and determine the relevant subject interests. A user survey reveals the frequency with which queries are put to the system, the form of response (i.e. reference or abstract), and the urgency of response, as well as the range in date of the sought information.

2.17 LIN, Nan and GARVEY, William D.: Information needs and uses. In: Annual review of information science and technology v.7, 1972.

2.18 VICKERY, B.C.: op.cit. p.17.

2.19 Ibid. p.48.

"Inquiries vary greatly in the amount of effort and time, and the type of skill needed to answer them satisfactorily, but the overall pattern is likely to remain fairly steady in any one library over long periods." 2.20

This observation can be seen later to be applicable when we approximate the data obtained from any monthly study of user needs to our results for one year. (see Figure III,3 p.64).

"In scientific and technical libraries one can expect that about one-fifth of the demand will be for simple facts and figures: questions which can often be answered readily from standard handbooks, encyclopaedias, data compilations and other reference works. About another quarter is likely to comprise essentially requests for a description of an object, a process, a method or a concept ..." 2.21

We have, however, found that the percentages given here by Vickery deviate from our findings after examination of inquiries handled by the Ciba-Geigy library. The first category of requests for simple facts corresponds to our Category 1 in which we reach a figure of 48% or almost exactly double that suggested by Vickery. The significance of this is that it shows that the regular needs of the staff for everyday reference - "to obtain specific items of information essential for the day-to-day conduct of an ongoing project" 2.22 - assumes far greater

---

2.20 Ibid. p.51.

2.21 Loc.cit.

2.22 VICKERY, B.C.: Information systems.
London, Butterworths, (©1973) p.2.

importance within this company in South Africa. It shows, in addition, that percentages and statistics applicable to a European special or industrial library cannot be extrapolated to fit our South African situation.

Vickery continues to point out that "it will be seen that nearly half of the subject inquiries in scientific and technical libraries are likely to make relatively modest demands on the skill and time of the inquiry staff". [2.23] Vickery presupposes here that user needs, fifty percent of the time, can be met by "one or two documents". In this library system, however, if this percentage were accepted, user needs would be badly underestimated. In Category 1, approximately half of the inquiries were answerable in this way: i.e. approximately 23% of the total number of inquiries. This interesting finding will be discussed in detail in Chapter 3.

This finding suggests a need for highly qualified information staff in order to satisfy user needs, with resultant high costs in providing the information required. It remains necessary, however, to evaluate these findings before the more detailed evaluation of the library system is discussed.

2.23 VICKERY, B.C.: Techniques of information retrieval. London, Butterworths, (1971). p. 52.

CHAPTER THREE

AN ANALYSIS OF THE DISSEMINATION OF INFORMATION FROM THE LIBRARY SYSTEM

3.1 THE USER SURVEY

This chapter examines user requests put to the library system during the period March 1973 to January 1974. In addition, a breakdown of requests reaching the library system via medical representatives and pharmaceutical division personnel indicates the wide geographical dissemination of information to the medical profession.

User requests were examined in order to attempt to establish patterns as regards level of complexity of inquiries, sources of information, level of user satisfaction and staff handling times for requests of varying complexity.

3.1.1 INTRODUCTION

The aim of any user survey is to establish the breadth and depth of information requirements in the subject fields covered and to discover how far the information centre fulfils current needs. If for example, it is discovered that a large percentage of the information required is obtainable from external sources only, then the user survey, by bringing fresh user requirements to light, may reveal such weaknesses.

3.1.2 THE OBJECTIVES OF THE USER SURVEY

The objectives of the user survey were to ascertain and measure the demands made on the system by its actual users.

3.1.3 METHODOLOGY

There are three methods which may be used in order to obtain information on user needs and user satisfaction. These are to collect the opinions of the users by means of:

(a) questionnaires

(b) interviews

(c) diaries as the user undertakes his search.

All of these methods are subjective and open to bias. The second method was chosen for this survey as it was felt that there would be fewer omissions in an interview. The third method places a burden on the user to record search steps taken and, in a busy period, users might omit comments as their searches proceeded. The second method was quicker and more accurate as the same interviewer questioned each user during the survey period. Before commencement of the search, each user was asked to specify the search area, comprehensiveness and source material required. After the user had indicated satisfaction, the break-off point in the search was noted together with the documents or document surrogates desired. The level of satisfaction attained was estimated by the inquirer as high, fair or low.

3.1.4 INQUIRY SAMPLE

The inquiries submitted to the library system between March 1973 and February 1974 were analysed by regression analysis (see Chapter 3.2.3 p.62). These were records of the actual use of the system from the library management point of view.

The following information was gleaned from an analysis of these inquiries.

(a) the distribution within and without the company of the inquiries.

(b) the type of answer supplied.

(c) the source(s) of the information provided.

The analysis of inquiry answers by source indicated the frequency of consultation of library materials (books, journals, product information in photocopy, microfilm or microfiche form; internal reports; and correspondence). (see Chapter 3.1.7 p.47).

The type of answer supplied has been categorized according to level of comple y (see Chapter 3.2 p.57; Figures III.2 p.63 and Figure III,3 p.64).

An analysis of the distribution of inquiries from within the company was made. (see Chapter 3.1.5 below). Figure III,1 p.44 shows the diffusion of information geographically to persons and organizations outside the company in Southern Africa.

D.V. Arnold has stated that "the larger the company, the more difficult it becomes to obtain up-to-date knowledge of people's

detailed current interests and the more valuable, as a source of information, became the questions and requests for service that are put to the various parts of the service." [3.1]

These requests, therefore, formed the basis of the survey which examined 798 inquiries submitted to the system.

3.1.5 DISSEMINATION OF INFORMATION

Because of the information flow to persons predominantly outside of the company itself (see Figure III,1 p.44), it was decided that the only way to undertake a survey was to examine actual inquiries which reached the company and which company staff referred to the library. In this way, inquiries from the medical profession via the company's medical representatives were also included within the survey.

The British AIOPI [3.2] symposium on 'Information needs of the prescribing physician' brought to light the fact that "the majority of general practitioners look to medical representatives as their principle source of drug information." [3.3] It is for this reason that medical representatives' requests for information for doctors in their sales areas were included in the user survey. It is interesting to note that inquiries originated from all parts of the country and

---

3.1 ARNOLD, D.V.: An industrial information service: the effects of growth on its organization and administration. In: Aslib proceedings 16(8): 240, 1964.

3.2 Association of Information Officers in the Pharmaceutical Industry.

3.3 Whose responsibility is drug information? In: Scrip 229: 10, 30 October, 1976.

that information from this library was diffused over a wide area. The map of South Africa shows this diffusion of drug information during the period February 1973 to the end of December 1973.

(see Figure III,1 overleaf)

The figures in brackets following the names of cities and provinces on the map are the total numbers of these inquiries in answer to which the company's library provided information. Smaller towns such as Ermelo, Vanderbijlpark and Carletonville are not marked, but the number of single inquiries originating from small towns within each province is totalled and placed in parentheses next to the name of the province, to provide a general indication within Southern Africa of the geographical origins of inquiries put to the library system.

There were 47 inquiries from doctors in Natal, via medical representatives: 153 inquiries from the Transvaal, 10 from the Orange Free State, 38 from the Cape, 3 from South West Africa and 8 from Rhodesia.

In addition to these, there were single inquiries directly from each of the following countries - France, Brazil, Uruguay, Mocambique, Czechoslovakia, Japan, Austria, Holland, Sweden and Denmark. There were two inquiries from each of the following countries - Canada, Germany, Thailand and Mexico; four from Switzerland; thirteen from the United States of America; seven from the United Kingdom; and three from the Far East. The library also supplied information to company headquarters in Basle - 22 items of information were forwarded.

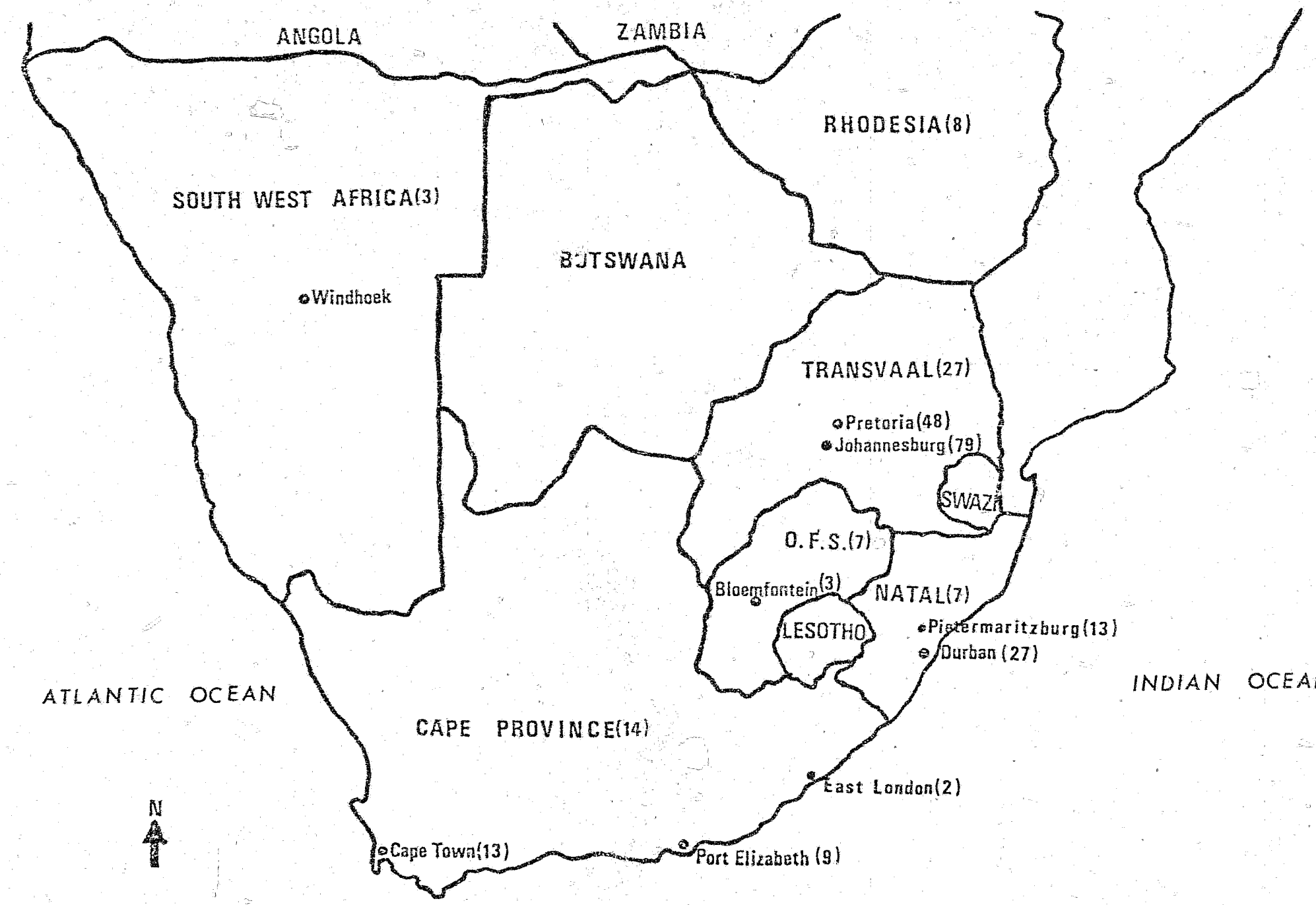


Figure III.1 Diffusion of Information throughout Southern Africa

Excluded from these figures are inquiries which originated within the South African company itself and it was usually not possible because of the inter-relationship between departments in the company, to analyse these inquiries according to the object of the inquiry as Cole [3.4] did in the British Petroleum Company's Technical Information and Library Service.

The inquiries which originated within the company, were included for analysis in Chapter 3.2. The breakdown of inquiries into three categories is made according to level of complexity of the inquiry and source of information. It is felt that these overall statistics would provide a more valid basis for system evaluation than a study within the company of the departmental origins of inquiries, as inquiries are frequently passed on inter-departmentally before reaching the library.

3.1.6. INFORMATION NEEDS

From the investigation of Ciba-Geigy library's users,* it became apparent that three types of information were sought. Firstly, there was an information need to verify what was already known or had been read (for example, by professionally trained people), leading to requests for standard textbooks or handbooks of pharmacology, medicine or even technical and language dictionaries. Included also were requests for information in which complete or almost complete bibliographical citations were given and all that was required of the

---

3.4 COLE, P.F.: The analysis of reference question records as a guide to the information requirements of scientists. In: Journal of Documentation 14(4): 197 - 207, 1958.

* They comprised: chiefly, medical practitioners and specialists in hospitals and private practice throughout South Africa; then medical practitioners, marketing and biostatistical staff employed by the company; and lastly institutional and retail pharmacists and nurses.

library staff was to extract and photocopy the desired information. Many of these were requests for journal articles which had already been read i.e., the full contents of which were known. Detailed examples will follow later.

Secondly, there exists a more important user need for information to solve an everyday problem, information which is required as soon as possible on a new competitive product or a new drug combination - linked usually to the corporate Marketing function.

Thirdly, there exists the need for information in depth. This may be defined, in this context, as the need for a comprehensive, detailed investigation of a topic, with nothing omitted i.e. total recall from within the system, and with the inquirer given access to all information, not just a carefully selected relevant portion of it (as in Category 2).

Examples include: all possible publications by a medical specialist who could possibly become a trialist for a locally pre-marketed drug; or all pre-clinical publications on that drug as background information for a clinical research trial; all possible publications on a new indication for a drug as background reading for the formulation of the methodology for a local clinical trial.

The types of information requirement which were met by the library within the year period of the user study can be seen in Figure III,2. p.60). Most of the internal inquiries came from three areas within the company, i.e. Management and Administration; Sales and Marketing; and Medical. These inquiries originating externally were sometimes directed to the library initially, but were usually

received chiefly via the Sales, Marketing and Medical departments. The Production function of the Company provided few inquiries for the library and it must be presumed that most of its needs were met by circulating journals in the industrial chemistry and pharmacy subject fields or that the library was not utilized.

Because of the number of information requests from medical practitioners outside the company itself, one conclusion reached was that the medical profession has a definite need for information on pharmacology and therapeutics which is not routed to conventional medical libraries. This is possibly related to the greater subject specialization of this library in these fields (see also Chapter 7).

3.1.7 SOURCES WHICH SUPPLY THE INFORMATION REQUIREMENTS

Information needs were grouped into categories according to the sources which supplied the required information:-

(a) It will be seen, therefore, that Category 1 includes all that information which was provided from books or periodicals or other published material such as pamphlets, conference abstracts, etc., in which the author or exact source was known. It includes, in addition, information from technical and language dictionaries and textbooks and handbooks of medicine, pharmacology etc.

In general, the search path followed for all these requests was based on a scanning of the following tools: firstly, the in-house South African card catalogue index, then the in-house Basle computer-printed indexes. Additional search paths were very rarely followed as these tools invariably provided sufficient citations. The exceptions were requests to Basle. (see p.53).

Examples:-

Request: i) Obtain a copy of the article by Balestieri on maprotiline.

Source: BALESTRIERI, A.:
Clinical comparative evaluation of maprotiline, a new antidepressant drug: a multicentre study.
In: International pharmacopsychiatry 6(4): 236 - 248, 1971.

Request: ii) Provide a copy of the Trasicor article by Darsinos.

Source: DARSINOS, J.:
The response of thyrotoxic tachycardia to Trasicor, a new beta-adrenergic agent.
In: Current therapeutic research 10(12): 626 - 630, 1968.

Request: iii) Provide a copy of the article by Rompel on carbamazepine in the aetiology of migraine.

Source: ROMPEL, H. and BAUERMEISTER, P.:
Aetiology of migraine and prevention with carbamazepine (Tegretol): results of a double-blind cross-over study.
In: South African medical Journal 44(4): 75 - 80, 1970.

Request: iv) Provide a copy of the latest Quarterly report of the S.A. Reserve Bank.

Source: SOUTH African Reserve Bank quarterly Bulletin.

Request: v) What is the meaning of the word "prophylaxis".

Source: STEDMAN, T.L.:
Stedman's medical dictionary: a vocabulary of medicine and its allied sciences, with pronunications and derivations. 23rd ed., completely revised... Baltimore, Williams & Wilkins, ©1976, p. 1148.

Request: vi) Provide a copy of the Millichap article on Ritalin.

Source: MILLICHAP, J. Gordon:
Drugs in the management of mimimal brain dysfunction.
In: Annals of the New York Academy of Science 205(2): 321-334, 1973.

Request: vii) Obtain a copy of the recent symposium on new perspectives in beta-blockade.

Source: International symposium on New perspectives in beta-blockade; Scanticon, Denmark, 1972; edited by D.M. Burley... Horsham, Ciba, 1973.

Request: viii) Provide a copy of the paper in Drugs 9(3): 178 - 226, 1975.

Source: DAVIES, D.L. and WILSON, G.M.:
Diuretics: mechanism of action and clinical application.

Request: xiv) Provide a copy of the article by H.E. Reef on temporal lobe epilepsy.

Source: REEF, H.E.:
Temporal lobe epilepsy.
In: Leech 40(3): 63 - 65, 1970.

(b) Category 2 includes sources from books, periodicals, other published material such as abstracts and pamphlets, as well as microfiche and microfilm. In this category an element of personal selection by library staff plays a large part in the interpretation of user needs and in user guidance towards meeting user requirements.

Examples:-

Request: i) Are there any papers on Tofranil, not Tegretol, in phantomb limb pain?

Sources: BEAUMONT, G.:
The use of psychotropic drugs in other painful conditions.
In: Journal of international medical research 4 (suppl. 2): 56 - 57, 1976.

POVACZ, F.:

Tofranil treatment of severe painful conditions after injuries.

In: Wien. med. wschr. 124(25 - 26): 432 - 433, 1974.

SAWAMURA, S. and TANAKA, T.:

Therapeutic result of imipramine (Tofranil) in phantom limb pain following amputation.

In: Orthopedic surgery 16: 623 - 637, 1965.

[Card catalogue:- Tegretol - pain, phantom limb]

Request: ii) Can Butazolidin be used in the treatment of dysmenorrhoea?

Source: MacQUEEN, R.C.:

Phenylbutazone in the treatment of dysmenorrhoea.

In: Practitioner 154 (1104): 787 - 789, 1970.

[Card catalogue:- Butazolidin - dysmenorrhoea]

Request:- iii) Information on the pharmacology of beta-adrenergic receptor blockers.

GOTH, Andres:

Medical pharmacology: principles and concepts. 7th ed. Saint Louis, Mosby, 1974. p. 154 - 158.

Request: iv) Can Anturan be administered to a patient undergoing haemodialysis?

Source: HAMPERS, Constantine L.: [and others]

Long-term haemodialysis: the management of the patient with chronic renal failure. 2nd ed. New York, Grune and Stratton, (1973).

Request: v) Obtain information on the use of Anturan in venous thrombosis.

Source: STEELE, Peter P.: [and others]
Platelet survival and adhesiveness in recurrent venous thrombosis.
In: New England Journal of medicine 288(22): 1148 - 1152, 1973.
[Paper indexed in card catalogue under product, Anturan]

Request: vi) Has Anafranil been used in the treatment of cataplexy?

Sources: GUILLEMINAULT, Christian: [and others]
A study on cataplexy.
In: Archives of neurology 31(4): 255 - 261, 1974.

NARCOLEPSY and cataplexy.
In: Lancet 1(7911): 84, 1975.

NARCOLEPSY
In: British medical Journal 1(5956): 476 - 477, 1975.

TREATMENT for cataplexy.
In: British medical Journal 1(5952): 233 - 234, 1975.

SIGWALD, J.:
Therapeutic non-psychiatric indications for clomipramine: Parkinson's disease, narcolepsy, pain, angina, enuresis.
In: Congres de psychiatric et de neurologie de langue francaise, Tunis, 1972. Proceedings p. 44 - 45.
[Papers indexed Anafranil - cataplexy in Card catalogue.]

Request: vii) Is rifampicin used for urogenital and renal tuberculosis?

Sources: HORNE, N.W.:
Renal tuberculosis.
In: South African Journal of hospital medicine 2(1): 16 +, 1976.

Sources: WECHSLER, Michael:

An evaluation of the current therapeutic regimen for renal tuberculosis.

In: Journal of urology 113: 760 - 761, 1975.

BELL, G.M. and DE KLERK, J.N.:

Tuberkulose van die penis (corpora cavernosa).

In: South African medical Journal 50(38): 1489 - 1490, 1976.

[Card catalogue:- Rimactane - tuberculosis; renal; tuberculosis, urethra]

Request: xiii) Obtain publications on the use of imipramine in the treatment of arthritis.

Sources: SCOTT, W.A. McDonald:

Imipramine aids arthritis pain.

In: Drug trade news 45: 52, 1970.

SCOTT, W.A. McDonald:

The relief of pain with an antidepressant in arthritis.

In: Pratitioner 202(1212): 802 - 807, 1969.

SCOTT, W.A. McDonald:

The use of psychotropics in rheumatic disease.

In: Rheumatologie 3(3): 303 - 206, 1973.

GRINGRAS, M.:

A clinical trial of Tofranil in rheumatic pain in general practice.

In: Journal of international medical research 4(suppl.2): 41 - 49, 1976.

Sources: HART, F. Dudley:

The use of psychotropic drugs in rheumatology.

In: Journal of international medical research 4(suppl.2): 15 - 19, 1976.

[Card catalogue:- Tofranil - arthritis, rheumatoid]

(c) Category 3 includes intensive searching through all available sources - as in Category 2 - but incorporates, in addition, verification by subject specialists such as Pharmacists, Medical Practitioners, in consultation with inquiry staff, in order to ensure that the information provided is comprehensive.

It involves, in addition, a search through in-house Computer-printed indexes which are the output from the company headquarters in Basle. Other relevant indexes and catalogues are also searched intensively and overseas sources of information are used exhaustively to provide as many as possible, if not all the documented works or papers on a specific topic.

Many requests to Basle fall into this category. If insufficient material is unearthed in the in-house search, the Documentation Centre in Basle is requested - often by telex - to provide a full bibliography on the subject. These requests may, therefore, serve as a check on the South African library's collection coverage and information retrieval. Should the bibliographical citations be known, and the relevant journals located in the Johannesburg area, these requests are then obtained locally through inter-library loans.

Examples:-

Request: i) Can Tegretol cause lupus erythematosus?

Sources: TAKIGAWA, Masahiro: [and others]

IgA deficiency and systemic lupus erythematosus.

In: Archives of dermatology 112(6): 845 - 849, 1976.

HARPEY, J.P.:

Drugs and 'disseminated lupus erythematosus'.

In: Adverse drug reaction bulletin 43: 104 - 143, 1973.

[Card catalogue:- Tegretol ADR dermatology]

Computer-printed indexes 1971 - 1977:-

SE/lupus erythematosus/systemic.

FROESCHER, W.

Nebenwirkungen antiepileptischer Medikamente.

In: Dtsch. med. wschr. 100(45): 2365 - 2366, 1975.

WIDMER, O. and KREBS, A.:

(Cutaneous side effects of internally administered drugs. Part iv of a synopsis. Drugs with effects on the central nervous system.)

In: Dermatologica 152(4): 193 - 260, 1976.

WEBER, M.:

(The choice of antiepileptic drugs based on their side effects.)

In: Cah. med. 15(1): 32 - 33, 1974.

Sources: LIVINGSTONE, S.: [and others]

Carbamazepine (Tegretol) in epilepsy. Nine year follow-up study with special emphasis on untoward reactions.

In: Diseases of the nervous system. 35(3): 103 - 107, 1974.

BEUREY, J.: [and others]

(Acute lupus erythematosus possibly induced by Tegretol)

In: Bull. soc. derm. syph. 79(2): 186, 1972.

Request: ii) Provide published counter arguments to the alleged association between rauwolfia alkaloids and breast cancer.

Sources: IMMICH, Herbert:

Rauwolfia derivatives and cancer.

In: Lancet 2(7883): 774 - 775, 1974.

SAXEN, E.A.:

Rauwolfia derivatives and breast cancer.

In: Lancet 2: 833, 1974.

BOCK, K.D.: [and others]

Rauwolfia derivatives and cancer.

In: Lancet 2(7885): 895, 1974.

RAUWOLFIA and breast cancer.

In: British medical Journal 4(5937): 121 - 122, 1974.

MANN, Ronald D.: [and others]

Rauwolfia derivatives and breast cancer.

In: Lancet 2(7886): 966, 1974.

MAGNUS, Knut:

Rauwolfia derivatives and breast cancer.

In: Lancet 2(7888): 1080, 1974.

HYPERTENSION 'strongly related' to cancer deaths.

In: Medical tribune 15(48): 2 - 3, Dec. 25, 1974.

GRIMSHAW, R.S. jr.:

Experts dubious on breast cancer tie with rauwolfia.

In: Medical tribune 15(40): 1, Oct. 23, 1974.

3.2 ANALYSIS OF INQUIRIES

3.2.1. INTRODUCTION

From the user survey a pattern of inquiries emerged. This can be followed on the graph which plots the number of queries put to the library system over a twelve month period, running from March 1973 to February 1974. It is interesting to note that in months in which a large number of inquiries were handled, the graph shows that the increase was fairly evenly spread amongst the three categories into which the queries were grouped. (see Figure III,3 p.64).

As Vickery [3.5] has stated that the overall pattern of inquiries is likely to remain fairly steady in a library over long periods it was necessary to establish strict criteria for grouping the inquiries according to time involved in finding relevant answers as well as the amount of staff skill required to find the required information.

3.2.2 ANALYSIS

In this particular scientific and technical library 47,86% of the inquiries were for simple facts or information answerable from dictionaries, encyclopaedias and other ready reference works, or to obtain a single printed or documentary proof or verification of what was already known, leading to requests for standard textbooks of pharmacology, medicine or technical and language dictionaries. An example is a definition of 'intractable pain' which required the use of textbooks of both medicine and therapeutics. The library's current awareness service provided the inspiration for further inquiries which were grouped into this category e.g.'plasma renin activity in hypertension and the use of beta-adrenergic receptor blocking drugs'was requested after articles on this subject has been listed in the monthly library current awareness circular

39,09% of the years' inquiries fell into Category 2 which comprised more complicated queries involving more staff time and requiring some degree of interpretation of user needs in satisfying user demand. The sources used for answers in this category were wider, and not

---

3.5 VICKERY, B.C.: Techniques of information retrieval. London, Butterworths, (1970). p.51.

confined to the ready reference material as in Category 1, and included abstracts, pamphlets, periodical articles as well as microfiche and microfilm company product literature. In this case the inquirer sought one or more relevant documents, but certainly did not require 100% recall of relevant documents in the system. Search time, therefore, was not considerable and there was the likelihood that the first few 'hits' would serve the user's needs.

In this instance, the grouping of 39,09% of inquiries into this category indicated that relatively modest demands were being made upon the library and upon the skill of the inquiry staff. Together with Category 1, the total percentage of inquiries was 86,95% which did not make heavy demands upon either time or skill of the library staff.

These findings suggest either that potential users were satisfied or were obtaining information from other sources e.g. outside libraries or learned colleagues, or else that this library's grouping of inquiries was different to that adopted by Vickery in obtaining his approximate total of 50%, or that the British data cannot be extrapolated to fit South African conditions, or that this library is not typical of South African or British scientific and technical libraries.

Cole's study [3.6] listed six types of answer supplied to requests. Grouped in a like manner to those of this study, they provide the following percentages.

Type 1	20%
Type 2	69%
Type 3	9%

with other types given as 2%

Compared with the percentages in this study there is no similarity, but

---

3.6 COLE, P.F.: op.cit. p.202

it is interesting to note that the total of Types 1 and 2 (89%) and Categories 1 and 2 (87%) shows a similarity.

In this survey, library loan records as well as references questions and answers, and all information searches were studied. All inquiries were entered according to date by library staff and these entries formed the basis for the analysis of queries put to the library system. It is, therefore, impossible to estimate now many people took their inquiries elsewhere as the library serves not only company staff, but also medical practitioners, pharmacists, pharmacologists, other paramedical personnel and biochemical laboratories.

The ideal of any library service is to provide the user with just the information he needs at the time ne needs it and in the form he desires. It is, therefore, a continual aim to improve performance not necessarily by providing for all possible needs, but by obtaining access to information that may meet possible needs through inter-library loans or, selective dissemination of information (SDI) services such as SASDI in peripheral subject fields. The library service must offer maximum availability of information in various forms and effectiveness can be improved by minimizing information loss to the user and by minimizing irrelevant recall. (see recall and precision ratios in Chapter 4).

In this user survey, it was found that there was a high success rate in the recall of relevant information (see Table III,1 p.61) but as almost 87% of the inquiries could be solved by the first correct 'hit', or the first few correct 'hits', user satisfaction alone cannot be the sole criterion for evaluating effectiveness.

It is, therefore, necessary to examine in greater detail those

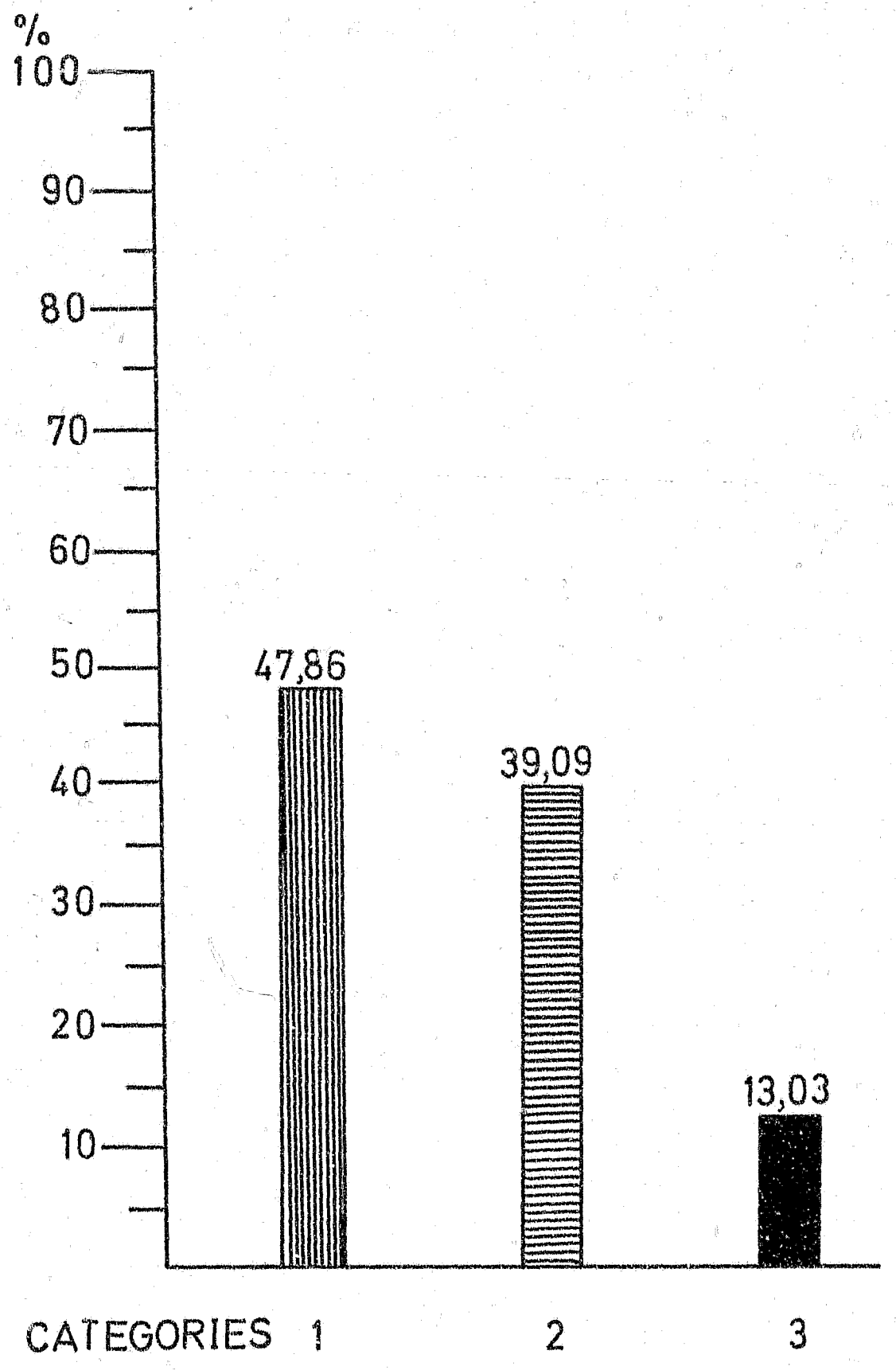


Figure III,2 Categorized Inquiries Expressed as a Percentage of the Year's Total

inquiries which were grouped into Category 3, in which total recall was required. These complex inquiries demanded wide subject knowledge and/or powers of judgement and competence in the library staff handling them. They comprised 13,03% of the total inquiries handled within the twelve-month period (see Figure III,2 p. 63).

In January 1975, inquiries for the month were studied and grouped, according to the same criteria, into three categories. The duration of time it took to provide satisfactory replies for the inquiries was noted. The breakdown is shown in Table III,1.

Category	Percentage of Total Inquiries	Average Time per Inquiry
1	60	5,0 mins.
2	30	11,6 mins.
3	10	27,5 mins.

Table III,1 Average Retrieval Time per Inquiry Category

The average time per inquiry enables one to calculate the average cost-in-time of answering each type of inquiry. These costs, however, are variables according to the salary levels of the staff answering the inquiries and do not take into account overheads. As this study was not designed to evaluate the system from a cost-effectiveness point of view, all that may be said of these average times is that they serve to emphasize the time factor involved in providing for the complexity and comprehensiveness requirements of producing search results in this environment.

It is interesting to note that there were approximately four

times the number of Category 1 queries as compared with Category 3 queries during the period analysed. The degree of selectivity needed for this type of inquiry in order to match documents to user needs was high and time consuming. Different indexes in use in the library system may have an effect upon the selection of documents, upon the recall and precision ratios, andit is for this reason that it was decided to formulate an experiment to evaluate these indexes (see Chapter 4).

The percentage of Category 1 inquiries satisfactorily answered was 99,55%; Category 2 was 98,09%; and Category 3 was 98,08%; which averages 98,57%.(see Table III,2 p.63)This is a very high average and it must be presumed that, of the users who expressed satisfaction at the time of the search interview, a number must later have reconsidered their judgements and returned to the library with modified search requests or else taken the modified requests elsewhere.

3.2.3 STATISTICAL ANALYSIS

Regression analysis showed that there is a significant increase in Category 1 queries over this twelve month period and a non-significant decrease over the year of Category 2 queries. There were approximately four times the number of Category 1 queries as Category 3 over this period and about 1,3 times as many Category 1 as Category 2 queries. There were 3 times as many Category 2 as Category 3 queries in the same period.

Category 1	S	Category 2	S	Category 3	S	Monthly percentages of inquiries per Category 1	2	3
18	100	17	100	9	100	40,9	38,6	20,5
2	100	20	100	4	100	7,7	76,9	15,4
17	94,1	24	100	10	100	33,3	47,1	19,6
23	100	35	97,1	10	100	33,8	51,5	14,7
21	100	33	96,6	4	100	36,2	56,9	6,9
56	100	37	100	14	100	52,3	34,5	13,2
36	100	28	100	5	100	52,2	40,6	7,2
23	100	29	100	11	100	36,5	46,0	17,5
43	100	32	100	12	100	49,4	36,8	13,8
37	100	13	100	2	100	71,2	25,0	3,8
50	100	17	94,1	8	100	66,7	22,6	10,7
56	100	27	96,3	15	100	57,1	27,6	15,3
24	100	12	91,7	4	75	60,0	30,0	10,0
Mean	99,55		98,09		98,08			

S = percentage of inquiries satisfactorily answered.

$$\frac{S1 + S2 + S3}{3} = 98,57\ \%$$

Table III,2 Percentage User Satisfaction By Category

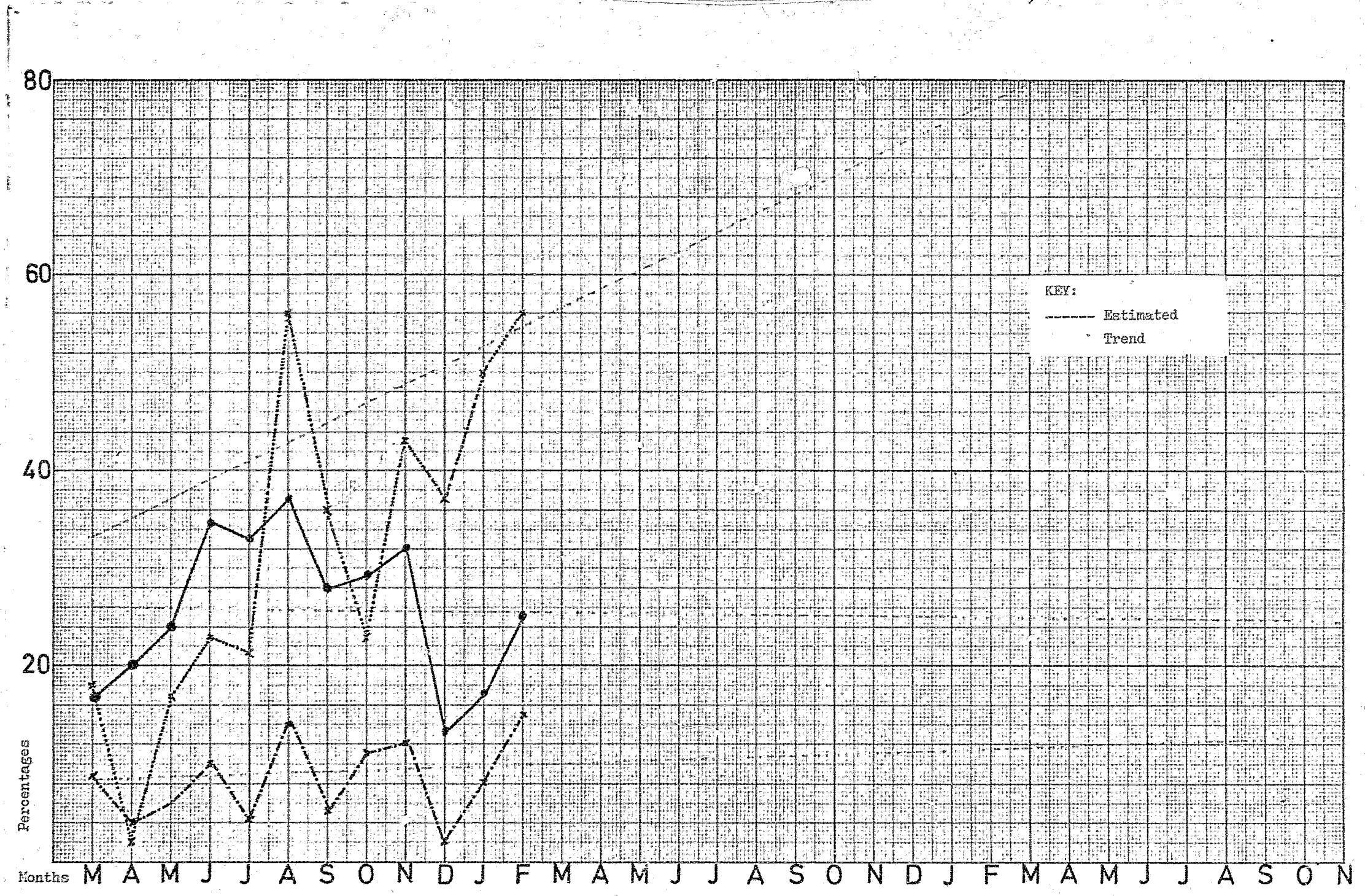


Figure III,3 Breakdown of Categorized Inquiries into Monthly Percentages: 1973 - 1974.

CHAPTER FOUR

LIBRARY SYSTEM EVALUATION

4.1 LIBRARY OBJECTIVES AND SYSTEM MEASUREMENT CRITERIA

4.1.1 INTRODUCTION

Library evaluation criteria are examined in relation to the operating effectiveness of the system, by considering its users, its objectives and its indexes. In other words, user satisfaction, system efficiency and effectiveness are its goals.

4.1.2 LIBRARY SYSTEM OBJECTIVES

Certain general library objectives need to be clearly demarcated. These include: the purchase of library materials, book, journals, reports and monographs (see page 67) in specific, relevant subject fields; the efficient distribution of these materials, especially periodicals and current awareness bulletins: the physical organization of the material to improve its availability; the successful communication between the user and library staff offering significant relevant information; and the selective dissemination of current information according to user interest profiles.

Public library objectives such as encouraging reading and cultural development and backing up community social activities fall away because of altered user requirements in an industrial environment. The potential special library user has often obtained a higher educational level than the public library user and special library users may even consist

exclusively of professionally trained persons. Objectives for the special library should, therefore, be defined in terms of company needs and the requirments of personnel involved in meeting those corporate objectives.

This statment of objectives must be detailed in order to yield criteria for evaluating alternative systems. It must also be used in order to determine workable library objectives which will enable library management to measure and evaluate library performance in meeting set objectives.

What aspects of a library system are measureable? The bookstock, journal subscription and issue statistics certainly are measureable, but are they an indication of the attainment of an objective such as user satisfaction? Measurement of library materials will indicate the growth of the library (see Chapter 7, p. 153 - 154).

4.1.3. THE DISSEMINATION OF CURRENT INFORMATION

The objective of dissemination of information in industry must be examined i.e., the transfer of knowledge from the originator to the user who absorbs, comprehends and stores it, and who, in a complicated process, may use it to convert it into new ideas and maybe even new publications.

The complicated dissemination process may be studied by means of user surveys. As the whole picture may vary if broken down week by week, month by month, or year by year, each of the following criteria must be examined in turn:- the collection coverage, the selection of new material, indexing, the current awareness service format, the retrieval of current material and the interaction of information and user.

The first criterion, collection coverage, is regarded as comprehensive in this library. This is because the major journals in the main subject fields are subscribed to, as well as others in peripheral subjects; the activities of group companies world-wide ensure that all relevant published papers in clinical medicine, pharmacology and toxicology in their geographical areas are covered and reach the company's Documentation Centre whereupon they are dispersed to other group companies.

The second criterion, selection of new material, is partially covered by group companies, as described above, and the selection is made as follows: All published papers as well as unpublished conference proceedings are covered in the country or territories which the group company serves (the South African company covers South Africa, Rhodesia, Botswana, Swaziland and South West Africa). Any mention of a company product either by trade name or by generic name ensures that the document will be included in the system. Certain value judgements are, however, made by the Documentation Centre and these determine whether the document will be classed as a 'mention' article or will be microfilmed and distributed to all group companies. There is, nevertheless, access to all these documents via the Documentation Centre itself. Locally, the company library selects documents on the same principles for inclusion in the system in photocopy form. After scanning of journals, those selected articles are indexed and, if judged to be of considerable importance, immediately despatched to the individual concerned in the form of photocopies of entire documents. This selection for dissemination requires a knowledge and appreciation of the precise needs of the small circle of users in the company's pharmaceutical division.

The third criterion, indexing, is linked to the fourth, current awareness service format. The current awareness circular contains new items of information (document surrogates) grouped together in a particular way. In this case, according to company product. If there is a weakness in this indexing, or if the format of the circular presents a barrier to easy scanning, then the dissemination of current information will be impaired. The provision of an SDI service can increase the efficacy of the dissemination of information from the system by facilitating an individual's access to new information as soon as possible after the publication and receipt of that information.

The fifth criterion, retrieval of current material from the library system, is as important as the two previous criteria. If the system's mechanical aids are overloaded and the user cannot obtain a copy of the desired document immediately he requests it, then his regard for the effectiveness of the library's performance will be lowered.

If the receiver of the disseminated information cannot interpret or use the information supplied, if it does not sometimes result in new ideas, the whole dissemination process is wasted. If it were possible to measure this criterion, then much of the experimentation on effectiveness of information retrieval might be channelled into finding ways to increase user usability of newly published information.

In order to maximize the dissemination of current information, all these factors need to be interlinked, and those measureable criteria need to be analysed so as to optimize library effectiveness.

4.1.4 LIBRARY SYSTEM MEASUREMENT CRITERIA

As this is a study of an operational system, it is necessary to consider the suitability of applying various measurement criteria in order to evaluate the performance of the system. Library effectiveness is measured to examining the following factors:-

a) Accessibility

b) User satisfaction

c) Response time

d) Use

e) Cost-benefit ratio

f) Cost

Accessibility to the library and its document store is a difficult criterion to measure. Direct user access to this library is restricted to company personnel and the occasional visitor. Indirect access by means of requests by letter, telephone and telex have been taken into account in Chapter 3.1.5. This aspect is quantifiable, but the second aspect of accessibility - the ease of determining the existence of documents in the document store and their location - is extremely difficult to measure. Availablility of information on the spot and coverage must be taken into account. Thus, considering only a certain group of frequent, well-catered for users will not measure total library performance accurately. The needs of the occasional user with the unexpected request must also be taken into account. Accessibility must be examined together with user satisfaction and use.

User satisfaction as a measurement criterion has two aspects, satisfaction with present services and need for non-available services.

The former has been measured (see Chapter 3), but the latter requires the examination of needs that were not met. This was not possible because of the geographical distribution of the library's potential users. The measurement of user satisfaction is subjective, but is able to highlight shortfalls in the users' expectations of the service.

It is possible to measure how many individuals present inquiries to a given library system and by careful questioning to elicit their degree of satisfaction with the information provided. But is this criterion alone suitable for use in evaluating a library system? There are those who seek information from outside sources and from communication with scientfic or technical colleagues or institutions (as discussed in Chapter 1.1) and who would render the latter criterion incomplete, or at best subjective. Yet it remains necessary to elicit some actual users' opinions of the library system.

For practical purposes, and in order to eliminate the pitfalls of questionnaires, these users were questioned while they were making use of the library, or after the search had been completed. In most cases, in this library, the information staff undertake the search. This is because of the time-saving factor involved. The information need is elicited by careful questioning, then the request is transformed into the language of the system by the information staff and the inquirer is asked to indicate the extent of searching required. At any stage the inquirer can stop the search when he has obtained sufficient documents relevant to his need. (see Chapter 3.1.3 p. 40).

The problem encountered here was the validity of responses by the library's actual users during the search interview in which they expressed satisfaction with the search results i.e., the relevancy of

documents or document surrogates (see Chapter 3, Table III,2 p.63).

Because of the service function of a library, user satisfaction must be considered as one of the most important evaluative criteria of library effectiveness, but it must be used in conjunction with the criteria of use and response time.

Response time is a useful evaluative criterion which is easily measured (see the average response times for the categorized library inquiries in Chapter 3, Table III,1 p.61). The response time in this case reflects the time taken to retrieve a relevant document or documents and to copy the document(s). If the response time becomes too long from the user's viewpoint, then the system does not perform satisfactorily. But a system which satisfies with regard to response time may not satisfy in other respects, such as accessibility. Response time is not meaningful applied in isolation and should be applied in conjunction with other evaluative criteria in order to produce meaningful results. If the system performs well in this respect it may imply that the stock is small or that is is well arranged. Time taken to obtain material on inter-library loan, or, in this case, from the Pharma Documentation Centre in Ciba-Geigy, Basle, should be excluded from the reckoning.

Library use has long been adopted as an evaluating criterion of library system performance. This is the reason for book issue figures and journal circulation statistics measuring the number of items used. A rise in use is not, in itself, a measure of rise in performance as increased usage may be accounted for by an increase in user numbers. Use studies may, however, isolate unused material which could be weeded out to increase the probability of a more efficient collection from

the user's point of view. A problem in the employment of this criterion is that it disregards potential use, and it must be applied over a long period in order to remove variations in actual use patterns, which may be seasonal or fluctuate according to changing user interests. (see Figure III, 3 p.64).

Costs may be briefly mentioned as an evaluative criterion as it is difficult to place a value or cost figure on a service and on information. For example, a rough estimate of cost per inquiry per year, may be gained by dividing the 1973 - 1974 inquiry total, 798, by the library staff salaries and cost of new accessions for that year. This method ignores overhead costs.

Corporate cost awareness may tend to lower expenditure on books and journals, while increasing the cost-in-time of individuals seeking information outside the company library system. These costs are not easily measureable because of the difficulty of identifying this occurrence.

It is possible to identify three levels at which to evaluate a library system.

They are (a) effectiveness evaluation (at least, from the librarian's viewpoint)

(b) cost effectiveness evaluation

(c) cost-benefit evaluation

In this study, the approach is predominantly the measurement of the first level, effectiveness, which is, in effect, an evaluation of user satisfaction. Evaluation of user satisfaction should elucidate how well the service satisfies its users' needs. However, the frequent pitfall of such evaluation is the fact that it evaluates expressed needs and

ignores unexpressed needs. In other words, by determining user satisfaction one obtains quantitative data, for example, that the library is able to provide 80% of the documents requested by users immediately they are sought.

This method indicates, therefore, the usefulness of the library stock and provides some indication of the need for inter-library loans and co-operation with other libraries.

King and Bryant [4.1] have distinguished between "macroevaluation" and "microevaluation". Macroevaluation attempts to establish how well the system is performing, whereas microevaluation is concerned with the reasons for the performance results of the system and attempts to identify ways in which system performance can be improved.

Cost-effectiveness evaluation relates measures of effectiveness to cost. In other words, if the document delivery success rate of the system is 80%, what would be the cost of raising this capability to 85%. The analysis which determines the least expensive method of increasing the system effectiveness is a cost-effectiveness analysis.

A cost-benefit evaluation is the most difficult of the three levels of evaluation to perform, as it involves attempting to assign a monetary value to information. It relates the cost of providing an information service to the benefits of having the service.

4.1 KING, D.W. and BRYANT, E.C.: The evaluation of information services and products. Washington, Information resources press, 1971. p. 37.

The purpose of evaluation is to quantify the functions which the system performs and to measure the achievement of the objectives which are set out to be accomplished. System objectives need to be described, but may prove inadequate as a basis for evaluation e.g. cost reduction and increased accessibility may prove impossible to implement simultaneously, although both goals may have been laid down. Realistic practical goals need to be accurately verbalized.

Furthermore, it is necessary to distinguish between system goals and objectives, and evaluation objectives. In the system under discussion, the evaluation was based on the need to make decisions to modify an existing system and to monitor system effectiveness continuously. In this case, the evaluation was intended to be undertaken for the benefit of the system management, the company management, as well as the user population. By user population, in this sense, is meant those immediate potential users in company employ who fall under the company's pharmaceutical division.

It must be noted that an operational system was evaluated and this influenced the objectives of the evaluation to some extent, as it affected the decisions which depended upon the evaluation; and the status of the system. Funds, and therefore, staff are employed differently in establishing a system as opposed to operating an existing system. Evaluation implies quantification of the performance characteristics of the system and, with this in mind, aspects of measurement must be examined.

The measurement itself may be performed in two ways: namely, description and comparison. In this case, both have been utilized. Two

alternative systems which were both operational in the same environment - e.g. the pharmaceutical industry have been compared. It is precisely for this reason that Index Medicus has not been examined in detail as it was not designed specifically for use in the pharmaceutical industry. It has however, been evaluated from a comparative point of view in relation to indexing terminology and ease of access. (see Chapter 4 p. 85).

4.1.5 EFFECTIVENESS MEASU[illegible]

In order to establish performance effectiveness in information retrieval it is necessary to establish whether the information (document or document surrogate) is transferred by the active or passive mode.

Retrospective searching characterises the latter mode in which the user initiates the search. King and Bryant [4.2] subdivide this mode as follows:

1) the user has prior knowledge of a document's identity (by title, author, or subject matter). This corresponds roughly, to Category 1 search inquiries, with the exception of the identification by subject matter which would fall into Category 2.
2) the user requests desired references by subject, without prior knowledge of specific documents.
3) the user browses to generate new ideas or to investigate new fields.

In the first case the emphasis is on document location, while in

---

4.2 Ibid. p. 20.

the second, document identification is achieved by means of retrospective searching. Browsing suggests documents or document surrogate perusal.

Document identification in the active mode is achieved by action on the part of system operators, and document or document surrogate identification is a result of prior agreement between the user and the system operators, whereby the system disseminates identification information - in this case by author, product or subject.

In general, it is the document or document surrogate identification process which must be evaluated. The measures of the effectiveness of the identification process are speed, timeliness, accessibility and accuracy, which can be related to system benefits.

The number of searches, number of references identified and the search time are cost related measures.

King and Bryant emphasize the importance of correct document identification. "Particularly important from the standpoint of evaluation, is the search situation where users need references in a subject without prior knowledge of specific documents. This situation is important because in identifying correct documents it is subject to error. The ability of the system to identify correct documents without error may be called system accuracy." [4.3]

King and Bryant define relevance [4.4] as the relationship between a user's information question (or information requirement statement)

4.3 Ibid. p. 22

4.4 Loc. cit.

and a document. The user's relevance judgement expresses this relevance relationship as a dichotomy (the document does or does not answer the user's information question). The system's response to the users inquiry may also be a dichotomy (the document is identified or not identified) and "the system relevance response is the systems assessment of the relationship between the user's question and a document" [4.5]. They define accuracy as "the degree to which system relevance response predicts the user relevance judgement, and an interesting evaluative criterion emerges from their statement that "system relevance responses are independent of user relevance judgements" and "accuracy must not be considered the user relevance judgement or system relevance response alone, but rather the degree to which the system relevance responses resemble the user relevance judgements" [4.6]

Performance measures using these criteria are implemented in a sample of inquiries during the period May to September 1978, which is analysed in Chapter 5. A document identification process is evaluated by utilizing a user relevance judgement measure in the following section.

4.5 Ibid. p. 23.

4.6 Ibid. p. 24.

4.2 EVALUATION OF INDEXES IN USE IN THE LIBRARY SYSTEM

4.2.1 INTRODUCTION

In addition to the comparison of indexes at present in use in the library, an evaluation study of these indexes was undertaken. Shortcomings of two of the indexes in use within the company have been indicated in other sections (see Chapter 6)and because these are subjective opinions in many cases, it was found necessary to include an experimental evaluation of the indexes currently in use in the library system in order to verify the hypothesis that the Computer-printed index has low recall and precision ratios and consequently lowers the effectiveness of information retrieval from the system.

4.2.2 BACKGROUND

There have been a number of recent experimental studies on the evaluation of printed subject indexes, viz. those by Lancaster, Cleverdon, Boon, Conoway, Jahoda, Keen and Drage. Conoway undertook a laboratory investigation of printed subject indexes and the "coefficient of index usability". [4.2] He defines "index usability" as the effectiveness and efficiency with which an index can be consulted. A series of sixteen alphabetical subject indexes to periodical articles was prepared using the index generation capabilities of the TEXT-PAC computer programmes. Synthetically generated questions were developed. Four variables were examined in sixteen test indexes. These were cumulativeness, abbreviation, selfstandingness and typography and he

---

4.7 CONOWAY, C.W.: An experimental investigation of the influence of several index variables on index usability and a preliminary study towards a coefficient of index usability. Buffalo (N.Y.), School of information and library studies, 1974.

provided a mathematical definition of the coefficient of index usability.

Special test indexes remove unwanted variables but problems such as the comparability of the search strategies and searcher-percieved relevance, performance measures and index measures, as well as the recording of the search procedure still remain.

Other experimental studies of printed subject indexes include the EPSILON project (Evaluation of Printed Subject Indexes by Laboratory investigation) [4.3] and the evaluation by J.A. Boon of Psychological Abstracts. [4.4]

In this section, index language as it affects the measures of recall and precision is examined.

The recall ratio is the number of relevant documents retrieved per inquiry in proportion to the total number of relevant documents in the collection.

The precision ratio (relevance ratio) is the number of relevant documents retrieved, in proportion to the total number of documents retrieved.

In order to ensure that every relevant document is retrieved it is, theoretically at least, necessary to search every document in the

---

4.8 KEEN, E.M.: A retrieval comparison of six published indexes in the field of library and information science. In: Unesco bulletin for libraries 30(1): 26 - 36, 1976.

4.9 BOON, J.A.: Paper presented at the International study institute on design and evaluation of printed subject indexes. Aberystwyth. 1975.

collection. A large number of non-relevant documents would be searched in order to provide relatively few relevant ones and thus the precision would be low. The precision ratio is, therefore, an important evaluative measure of the efficiency of a retrieval system. For example, 40 documents are retrieved in a search and of these only 5 are relevant. The precision ratio will, therefore, be:

$$\frac{5}{40} = 12.5\ \%$$

The recall ratio is the more difficult evaluative measure to apply. It implies that for each search it is possible to know the total number of relevant documents in the collection. Yet, if the search has yielded 5 'hits' out of 10, and thus a recall ratio of:

$$\frac{5}{10} = 50\ \%$$

it presupposes that the searcher/librarian will later find the remaining 5 relevant documents and then be able to work out the recall ratio. As the search is carried out, the relevant documents or 'hits' are presumed to comprise all the relevant material and therefore a recall ratio of 100 % - total recall.

Relevance assessment is problematical - is it to be made from the viewpoint of the system operators or that of the user. In evaluating the indexes in use in the library system, relevance will be obtained if the indexes succeed in matching document surrogates to a stated request and the relevance judgement may be made by a subject specialist - in this case, the information workers in the library. The information staff made the relevance judgements in this test in order to obtain precision ratios from the test results.

The precision ratio, in order to be meaningful, must be coupled

with the recall ratio. As mentioned above, the only way to accurately determine the recall ratio is to examine the total collection. In this test, each question was answered by use of all three indexes and the relevant document surrogates were compared in order to fill in any omissions. The company Computer-printed index must, however, be regarded as providing comprehensive cover, because of the world-wide network of information coverage by group companies.

The aims of the test were to determine whether there was a defective user-index interaction or an indexing, or index language weakness.

4.2.3 METHODOLOGY

Librarianship students were used for this test as it was felt that people unfamiliar with indexes and indexing terminology would give a less accurate result. This also accounts for the small sample size as only fourth year Library Science or Higher Diploma in Librarianship students were eligible.

Firstly, nine of the students were subjected to a mental alertness test for graduates - drawn up by the National Institute for Personnel Research to measure logical reasoning - and grouped according to scores. (see Table IV,3 p.82). It was hoped to find a correlation between high scores on this test which indicate good powers of logical reasoning, therefore effective search strategies, and hence good recall and precision ratios. A control group of three students did not undergo this test. Secondly all students were in turn, given a difficult query to solve using each library index. All subjects were restricted to a half hour per query. The queries used in this experiment were drawn from those examined in Chapter 3 and were all actual inquiries put

previously to the library system.

Table IV,1 lists the actual inquiries used to test the students' handling of the indexes. The level of difficulty was comparable as these queries were all regarded as Category 3 inquiries. (see Chapter 3.2.2 p.61).

1.	The Pharmacology Department of a University Medical School requires information on the mode of action of a product (S) in influenza.
2.	Has a product (A) ever been used in the treatment of amoebiasis?
3.	What effect (effects) does drug (T) exert on the central nervous system?
4.	Is product (R) indicated for the treatment of leprosy?
5.	Can product (SE) be administered by intra-arterial injection for vascular disease?
6.	Is product (B) indicated for tuberculous meningitis?
7.	Can product (T) be used in the treatment of bradycardia?
8.	Is product (TA) indicated in the treatment of tuberculous meningitis?

Table IV,1 Actual Inquiries Used in the Index Experiment

The candidates were grouped in order to test the hypothesis that powers of logical reasoning may affect index use and hence recall and precision ratios. The questions in Table IV,1 were categorized under three indices A, B and C. Candidates answered questions using each index. Questions used with each index were homogeneous insofar as

subject matter and degree of complexity was concerned. These actual queries were grouped in such a way that, to the greatest extent possible, the level of difficulty of questions in use on each index was similar i.e. all questions had been previously categorized as Category 3 (see Chapter 3.2.2 p. 61). Each candidate answered different questions using each index.

Written search requests were supplied to the twelve students with instructions to enter the search commencement and concluding times for each question on the answer sheets. The students were advised that the search would be through three years' issues/entries for each inquiry. They were informed that the inquiries were actual inquiries already answered through the use of the same indexes and were requested to list relevant citations within the thirty minute time period allotted for each search. Citations which were considered to be possibly relevant were also to be listed in the space provided on the answer sheets. (see Appendix 1). After concluding each search, each student was asked to note any opinions or problems encountered in operating the index and to indicate whether he was satisfied that he had found relevant references. In this manner it was possible to test the students' responses to each index without direct learning influencing responses to the subsequent indexes and thus to reduce misleading bias in the results due to interpretation of subject matter into indexing language.

The candidates were grouped into threes and were assigned to each index as follows:

The Card subject catalogue = (A), Index Medicus = (B), and Computer-printed index = (C).

Candidate	R	Index	A	B	C
Candidate	S	Index	B	C	A
Candidate	T	Index	C	A	B
Candidate	U	Index	A	B	C
Candidate	V	Index	B	C	A
Candidate	W	Index	C	A	B
Candidate	X	Index	A	B	C
Candidate	Y	Index	B	C	A
Candidate	Z	Index	C	A	B
Control	A		A	B	C
Control	B		B	C	A
Control	C		C	A	B

Table IV,2 Random Allocation of Test Inquiries

4.2.4 TEST RESULTS

The scores on the National Institute for Personnel Research (N.I.P.R.) 'Mental Alertness Test for Graduates' were as follows:-

Candidate	R	35
Candidate	S	34
Candidate	T	31
Candidate	U	28
Candidate	V	29
Candidate	W	24
Candidate	X	28
Candidate	Y	28
Candidate	Z	19

Table IV,3 Candidates' Scores on N.I.P.R. Mental Alertness Test for Graduates

The recall and precision ratios were determined for each search undertaken on the three indexes in use in the library, and were as follows:-

		Recall Ratio	Precision Ratio
Candidate	R	0,0 %	0,0 %
Candidate	S	83,3 %	62,8 %
Candidate	T	0,0 %	0,0 %
Candidate	U	0,0 %	0,0 %
Candidate	V	0,0 %	0,0 %
Candidate	W	100,0 %	100,0 %
Candidate	X	16,6 %	50,0 %
Candidate	Y	83,3 %	83,3 %
Candidate	Z	83,3 %	62,5 %
Control	A	0,0 %	0,0 %
Control	B	14,2 %	14,3 %
Control	C	0,0 %	0,0 %
Average		23,3 %	31,1 %

Table IV,4 Candidates' Test Results using Index Medicus

Although the sample size was small, it provided interesting information on the three indexes. The performance of the dictionary Catalogue card index which was designed for this library proved to have the highest average recall ratio of 50,2 % and highest average precision ratio of 35,1 % (see Table IV, 6 p.84). The average recall and precision ratio for Index Medicus were 23,3 % and 31,1 % respectively (see Table IV,4). The Computer-output index from company headquarters showed a consistently low performance with average recall and precision ratios of 19,7 % and 14,9 % respectively (see Table IV,5 p. 84).

		Recall Ratio	Precision Ratio
Candidate	R	100,0 %	75,0 %
Candidate	S	20,0 %	17,2 %
Candidate	T	0,0 %	0,0 %
Candidate	U	0,0 %	0,0 %
Candidate	V	0,0 %	0,0 %
Candidate	W	50,0 %	50,0 %
Candidate	X	16,6 %	14,3 %
Candidate	Y	0,0 %	0,0 %
Candidate	Z	0,0 %	0,0 %
Control	A	0,0 %	0,0 %
Control	B	0,0 %	0,0 %
Control	C	50,0 %	23,0 %
Average		19,7 %	14,9 %

Table IV, 5 Candidates' Test Results using the Computer Index

		Recall Ratio	Precision Ratio
Candidate	R	100,0 %	50,0 %
Candidate	S	50,0 %	25,0 %
Candidate	T	33,3 %	17,6 %
Candidate	U	77,7 %	53,9 %
Candidate	V	0,0 %	0,0 %
Candidate	W	100,0 %	80,0 %
Candidate	X	44,4 %	33,3 %
Candidate	Y	50,0 %	33,3 %
Candidate	Z	50,0 %	50,0 %
Controls	A	22,2 %	18,2 %
Controls	B	0,0 %	0,0 %
Controls	C	75,0 %	60,0 %
Average		50,2 %	35,1 %

Table IV, 6 Candidates' Test Results using the Catalogue Card Index

In this test it was document surrogates which were retrieved, not the actual documents themselves. It is interesting to note that no correlation was found in this small sample between search strategies which were almost 'noise free' and the students scores on the National Institute for Personnel Research (NIPR) Mental Alertness Test which measures logical reasoning. In order to evaluate this aspect of the search strategy, it is necessary to examine the precision ratios achieved by the candidates.

	NIPR test scores	Index Medicus	Computer Index	Card Index
Candidate R	35	0,0 %	75,0 %	50,0 %
Candidate S	34	62,8 %	17,2 %	25,0 %
Candidate T	31	0,0 %	0,0 %	17,6 %
Candidate U	28	0,0 %	0,0 %	53,9 %
Candidate V	29	0,0 %	0,0 %	0,0 %
Candidate W	24	100,0 %	50,0 %	80,0 %
Candidate X	28	50,0 %	14,3 %	33,3 %
Candidate Y	28	83,3 %	0,0 %	33,[illegible] %
Candidate Z	19	62,5 %	0,0 %	50,0 %

Table IV. 7 Candidates'NIPR Test Scores and Precision Ratios on Three Indexes

Multiple linear regression was performed on this data. No significant correlation was found between the NIPR test scores and the candidates precision ratios on Index Medicus, the Computer index and the Catalogue card index. In addition, there was no significant correlation between precision ratios on the Computer and Card indexes; on Index Medicus and the Card index; or on the Computer index and Index Medicus.

The highest correlation was found between precision ratios on the

Computer and Card indexes, although this was not significant. Both these indexes were designed to operate as tools in the retrieval of company product information.

These results are not, however, an exact indication of the actual operating efficiency of the Computer indexes as the test subjects had more difficulty finding index entry terms than the information staff of the library. The test subjects commented on this index as follows:-

"Extremely difficult to understand " - Candidate V

"Extremely difficult to use" - Candidate X

"Find the index difficult to understand" - Candidate U

"Very difficult to use" - Candidate T

"Difficult to understand explanations of how to facilitate the searching procedure" - Candidate Z

From the candidates comments, it was deduced that there was a defective user-index interaction in using the Computer index, and an index language weakness. The latter is illustrated by the low recall and precision ratios found in the test results for this index.

CHAPTER FIVE

MICROEVALUATION OF THE RETROSPECTIVE SEARCHING PROCESS IN THE LIBRARY SYSTEM

5.1 INTRODUCTION

An overall assessment of the effectiveness of the Ciba-Geigy library system has been discussed previously in Chapters 3 and 4. Whilst such overall evaluation is essential, it is, nevertheless, useful to attempt to evaluate the individual components of the system. This type of microevaluation, as opposed to the former macroevaluation, can indicate the strengths and weaknesses of the components of the system, which is an essential prerequisite if the system is to be rationally improved. The purpose of this chapter is, therefore, to study in greater detail the most important components of the library system as implemented at Ciba-Geigy, South Africa.

The method of analysis used is that described by King and Bryant and is briefly described in the next paragraph. Subsequent sections of this chapter describe the data collection, the statistical analysis of the data and the deductions which can be made from the analysis.

5.2 THEORY

5.2.1 GENERAL

The various processes which constitute the activities of the library system include the following:

1) Indexing, cataloguing, and other input activities.

2) File organization.

3) Retrospective searching.

4) Current awareness and selective dissemination of information (S.D.I.) services.

5) Screening and other output activities.

Most of these processes, as implemented in Ciba-Geigy's library system have been described elsehere;

indexing in Chapter 2.3.2 p. 29 and Chapter 6 p.106-113, 120-134.

retrospective searching in Chapter 2.3.2 p. 28 - 29.

current awareness and S.D.I. services in Chapter 2.3.4 p. 34 and Chapter 4.1.3 p. 67 - 68,

screening in Chapter 2.3.4 p. 34 and Chapter 4.1.3 p. 67.

The topic of this chapter revolves around retrospective searching, which forms a core activity of the library service at Ciba-Geigy, South Africa.

The system begins to operate when a request, usually in verbalized form, is submitted.

The request is interpreted by a coder and encoded into the search language, sometimes after interaction between the library staff and the inquirer. The encoded query is matched against the indexes in use in the library. It is presumed, for the purpose of this evaluation, that the two indexes, the Catalogue card index and the Basle-produced Computer-printed index, adequately express the contents of the documents on file. The documents which pass the query-file matching test are screened for relevance, and those passing the test are submitted to the user.

This completes one iteration of the search and may be repeated as often as necessary.

A careful analysis of the above steps shows that errors can creep into the system in the following ways:

(1) The person who encodes the request may misinterpret it (coder interpretation error).

(2) The system language may be so restrictive that the request encoded into that language may not reflect the same content as the verbalized request.

(3) The file may contain indexing errors.

(4) The screener may make errors in selecting documents from the output which match the verbalized request". [5.1]

The primary measure of the system's performance is the relationship between system relevance response and user relevance judgement (see Chapter 4.1.5 p.77). However, a number of operations which require judgement decisions in relation to the document sought intercede between the initial and the final output. King and Bryant's systems diagram, which shows the system accuracy and the relationships that contribute to it, is reproduced overleaf as Figure V,1.

The circles along the periphery of the diagram show the processes set in motion by a user question. For example, an intermediary interprets the question, he translates that question into a system query, the system matches the query and indexed documents in terms available to the system, and an intermediary screens documents from the system output based on his interpretation of the user's question.

5.1 KING, D.W. and BRYANT, E.C.: op.cit. p.89 - 90.

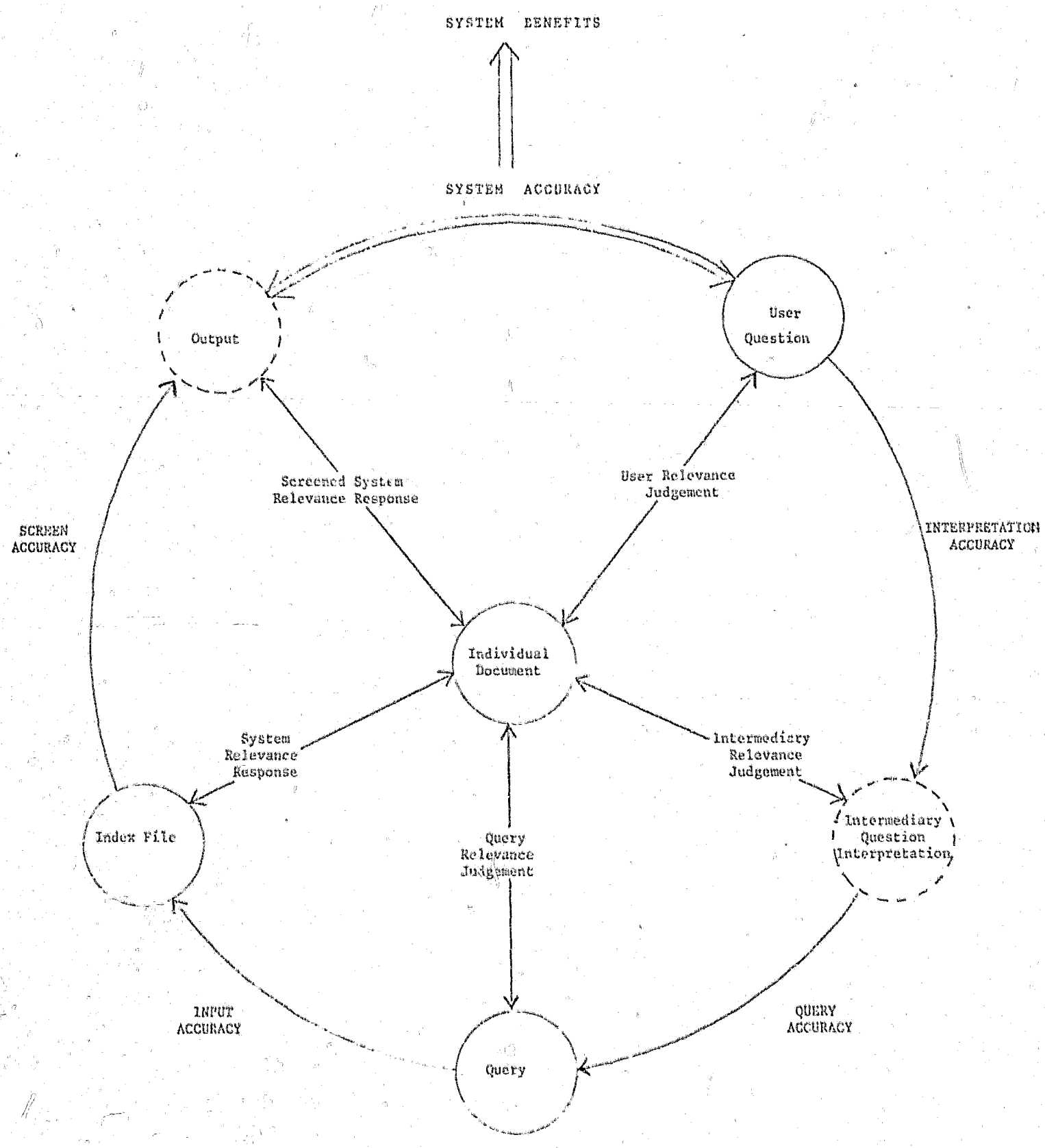


Figure V,1 Schema presenting System Accuracy and Relationships that contribute to it.

Each of these forms of the user question, shown in the peripheral circles, is related to documents in the file in a manner similar to user relevance judgement and should be highly correlated to user relevance judgment. Hence, starting at the upper right hand corner, the user relevance indicates the degree of relationship between the user question and a document. Similarly, progressing clockwise around the diagram, each radial arrow indicates a judgement, by the person concerned in that specific process, of the degree of relationship between that particular form of the query and an individual document. The five relationships can be multi-valued or dichotomous and the corresponding scores can be plotted against one another, as shown in Figure V,2. "Each point represents one document from the file. Each document has a user relevance judgement value and a value found by processing the user's question through the system. For a given user question, no documents in a file are likely to have low values on the scales of user relevance judgement and system relevance response, which accounts for the large cluster of

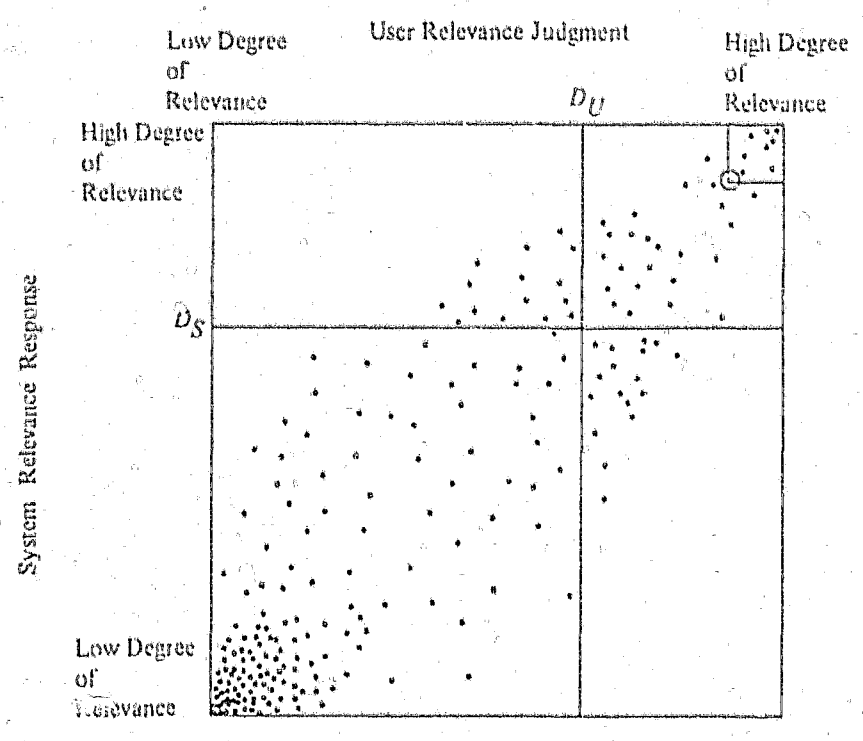


Figure V,2 User Relevance Judgement Plotted Against System Relevance Response for Individual Documents and a Given Search Question

documents in the lower left corner. It is highly unlikely that values of system relevance response will be the same as user relevance judgement. The degree to which system relevance response predicts the user relevance judgement is called accuracy". 5.2

Hence the relationships between the question forms and documents provide four accuracy measures, including interpretation accuracy, query accuracy, input accuracy and screening accuracy, as indicated in Figure V,1 p.92. A lack of accuracy in each case contributes to the deterioration of overall system accuracy and the relative contribution can be measured along each link in the chain. Each source of failure can be isolated and accuracy can be measured for interpretation, query, input, and screening. The problem is that all possibilities must be considered. For example, a relevant document may be missed due to interpretation error, but this document may be identified by the search query for the wrong reason. The only feasible way to account for all combinations of occurrences is by means of a mathematical model as proposed by King and Bryant and described below.

5.2.2 MATHEMATICAL MODEL

The mathematical model which describes the sources of error discussed above, makes use of conditional probabilities. In standard notation this is written P(A/B) which is read 'the probability of A given B'. The symbols used in the model have the following definitions:

V_r, relevant with respect to verbalized request;

$V_{\bar{r}}$, nonrelevant with respect to verbalized request;

5.2 KING, D.W. and BRYANT, E.C.: op.cit p.22 - 23.

C_r, relevant with respect to coder's interpretation;

$C_{\bar{r}}$, nonrelevant with respect to coder's interpretation;

E_r, relevant with respect to encoded request;

$E_{\bar{r}}$, nonrelevant with respect to encoded request;

R_r, relevant with respect to system's response;

$R_{\bar{r}}$, nonrelevant with respect to system's response;

S_r, relevant with respect to screener's judgement; and

$S_{\bar{r}}$, nonrelevant with respect to screener's judgement. 5.3

Thus $P(E_r/C_{\bar{r}})$ means 'the probability that a document is relevant to the encoded request, given that it is not relevant to the coder's interpretation'. In practice, one is working with relative frequencies, since the limiting values are unknown, but, for convenience, the conceptual limits are used and are called probabilities. These probabilities can be displayed in 2 x 2 tables as shown in Figure V,3 overleaf.

Perusal of this table shows which activities are the principal sources of error. Ideally, all the entries in these tables should be zeros and ones, with the ones in the lower left hand and upper right hand corners. The amount of departure from this ideal indicates the extent of departure from perfection. (see Chapter 5.5, p.105).

5.3 Ibid. p. 93.

	Relevance with respect to coder's interpretation	
Relevance with respect to verbalized request	$C_{\bar{r}}$	C_r
V_r	$P(C_{\bar{r}}/V_r)$	$P(C_r/V_r)$
$V_{\bar{r}}$	$P(C_{\bar{r}}/V_{\bar{r}})$	$P(C_r/V_{\bar{r}})$

	Relevance with respect to encoded request	
Relevance with respect to coder's interpretation	$E_{\bar{r}}$	E_r
C_r	$P(E_{\bar{r}}/C_r)$	$P(E_r/C_r)$
$C_{\bar{r}}$	$P(E_{\bar{r}}/C_{\bar{r}})$	$P(E_r/C_{\bar{r}})$

	Relevance with respect to response by system	
Relevance with respect to encoded request	$R_{\bar{r}}$	R_r
E_r	$P(R_{\bar{r}}/E_r)$	$P(R_r/E_r)$
$E_{\bar{r}}$	$P(R_{\bar{r}}/E_{\bar{r}})$	$P(R_r/E_{\bar{r}})$

	Relevance with respect to screener's interpretation	
Relevance with respect to verbalized request	$S_{\bar{r}}$	S_r
V_r	$P(S_{\bar{r}}/V_r)$	$P(S_r/V_r)$
$V_{\bar{r}}$	$P(S_{\bar{r}}/V_{\bar{r}})$	$P(S_r/V_{\bar{r}})$

Figure V.3 Conditional Probabilities Used in Retrospective Search Models

	Relevance with respect to coder's interpretation	
Relevance with respect to verbalized request	$C_{\bar{r}}$	C_r
V_r	$P(C_{\bar{r}}/V_r)$	$P(C_r/V_r)$
$V_{\bar{r}}$	$P(C_{\bar{r}}/V_{\bar{r}})$	$P(C_r/V_{\bar{r}})$

	Relevance with respect to encoded request	
Relevance with respect to coder's interpretation	$E_{\bar{r}}$	E_r
C_r	$P(E_{\bar{r}}/C_r)$	$P(E_r/C_r)$
$C_{\bar{r}}$	$P(E_{\bar{r}}/C_{\bar{r}})$	$P(E_r/C_{\bar{r}})$

	Relevance with respect to response by system	
Relevance with respect to encoded request	$R_{\bar{r}}$	R_r
E_r	$P(R_{\bar{r}}/E_r)$	$P(R_r/E_r)$
$E_{\bar{r}}$	$P(R_{\bar{r}}/E_{\bar{r}})$	$P(R_r/E_{\bar{r}})$

	Relevance with respect to screener's interpretation	
Relevance with respect to verbalized request	$S_{\bar{r}}$	S_r
V_r	$P(S_{\bar{r}}/V_r)$	$P(S_r/V_r)$
$V_{\bar{r}}$	$P(S_{\bar{r}}/V_{\bar{r}})$	$P(S_r/V_{\bar{r}})$

Figure V,3 Conditional Probabilities Used in Retrospective Search Models